



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 810 287 A1

(12)

# EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:  
03.12.1997 Bulletin 1997/49

(51) Int. Cl.<sup>6</sup>: **C12N 15/54**, C12N 9/10,  
C12N 1/21, C12N 5/00,  
A01H 5/00

(21) Application number: 96902462.9

(22) Date of filing: 16.02.1996

(86) International application number:  
PCT/JP96/00348

(87) International publication number:  
WO 96/25500 (22.08.1996 Gazette 1996/38)

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL  
PT SE

(30) Priority: 17.02.1995 JP 67159/95  
29.06.1995 JP 196915/95  
30.01.1996 JP 46534/96

(71) Applicant: SUNTORY LIMITED  
Kita-ku, Osaka-shi, Osaka-fu 530 (JP)

(72) Inventors:  
• ASHIKARI, Toshihiko  
Takatsuki-shi, Osaka 569 (JP)  
• TANAKA, Yoshikazu  
Otsu-shi, Shiga 520-02 (JP)  
• FUJIWARA, Hiroyuki  
Neyagawa-shi, Osaka 572 (JP)

• NAKAO, Masahiro  
Nagaokakyo-shi, Kyoto 617 (JP)  
• FUKUI, Yuko  
Mishima-gun, Osaka 618 (JP)  
• SAKAKIBARA, Keiko  
Takatsuki-shi, Osaka 569 (JP)  
• MIZUTANI, Masako  
Kyoto-shi, Kyoto 615 (JP)  
• KUSUMI, Takaaki  
Suita-shi, Osaka 564 (JP)

(74) Representative:  
Stoner, Gerard Patrick et al  
MEWBURN ELLIS  
York House  
23 Kingsway  
London WC2B 6HP (GB)

## (54) GENES CODING FOR PROTEINS HAVING ACYL TRANSFER ACTIVITY

(57) Proteins having an aromatic-acyltransferase activity; a gene system coding for the same; a process for producing the proteins using the gene system; and uses of the genes and the proteins. The genes and the proteins acylate plant pigments such as anthocyanin to thereby cause color tone changes, thus providing plants, particularly flowers, having colors that have not been inherent therein.

EP 0 810 287 A1

## Description

## Field of the Invention

The present invention relates to genes encoding proteins having aromatic acyl group transfer activity and the use thereof. More particularly, the present invention relates to genes encoding proteins having aromatic acyl group transfer activity derived from gentians (*Gentiana triflora* var. *japonica*), petunias (*Petunia hybrida*), perillas (*Perilla ocimoides*), and cinerarias (*Senecio cruentus*), and the use thereof.

## Background Art

The flower industry is making efforts to develop new and various varieties. An effective method of producing a new variety involves changing the color of a flower, for which the traditional breeding methods have been successfully employed to produce a wide variety of colors for almost all commercial varieties. With the above methods, however, it is rare that a single species produces colored varieties coming in a wide range of different colors since a pool of genes is limited for each species.

The colors of flowers are based mainly on two types of pigments, flavonoids and carotenoids. Flavonoids contribute mainly to the colors in the range of yellow to red and blue, while carotenoids contribute to the color tones of orange or yellow. Flavonoid molecules that make a major contribution to the color of flowers are anthocyanins which are glycosides of cyanidin, delphinidin, petunidin, peonidin, malvidin, and pelargonidin. Different anthocyanins impart marked changes in the color of flowers. Furthermore, the color of flowers is affected by copigmentation with colorless flavonoids, metal complex formation, glycosylation, acylation, methylation and pH of vacuoles (Forkman, Plant Breeding 106: 1, 1991).

There are a number of reports of acylated anthocyanins isolated from nature including cinerarin derived from cinerarias (*Senecio cruentus*) (Goto et al., Tetrahedron 25: 6021, 1984), awobanin derived from dayflowers (*Commelina communis*) (Goto and Kondo, Angew. Chem. Int. Ed. Engl. 30: 17, 1991) and gentiodelphin derived from *Gentiana Makinoi* (Yoshida et al., Tetrahedron 48: 4313, 1992) (Monarda didyma: Kondo et al., Tetrahedron 26: 5879, 1985; perillas, pansies (Goto et al., Tetrahedron 27: 2413, 1987; Wandering Jew: Idaka et al., Tetrahedron 28: 1901, 1987; *Dioscorea japonica*: Shoyama et al., Phytochemistry 29: 2999, 1990; red cabbage, *Platycodon grandiflorum*, lobelia, delphiniums, butterfly peas: Goto and Kondo, Angew. Chem. Int. Ed. Engl. 30: 17, 1991; carrots: Glabgen et al., Phytochemistry 31: 1593, 1992; morning glory: Lu et al., Phytochemistry 32: 659, 1992; Saito et al., Phytochemistry 40: 1283, 1995; *Alga decumbens*, *Clinopodium gracile*, *Lamium*, lavender, catnip, *Leonurus macranthus*, *Plectranthus*, *Prunella*, *Salvia splendens* Sella, Japanese Artichoke: Saito and Harborne, Phytochemistry 31: 3009, 1992; giant water lily: Strack et al., Phytochemistry 31: 989, 1992; bellflowers: Brandt et al., 33: 209, 1993; gentians: Hosokawa et al., Phytochemistry 40: 941, 1995; hyacinth: Hosokawa et al., Phytochemistry 40: 567, 1995).

Acyl groups which modify these anthocyanin-containing flavonoids are divided into two classes based on their structure: one is the aromatic acyl groups centering on hydroxy cinnamic acids, and the other is the aliphatic acyl groups such as the malonyl group. It has been observed in the experiment carried out using the anthocyanin pigment of morning glories (*Pharbitis nil*) that among the acyl groups transfer reactions anthocyanins to which an aromatic acyl group, preferably coumaric acid or caffeic acid, is bound show a shift of the absorption maximum to the long wavelength side (Dangle et al., Phytochemistry 34: 1119, 1993).

Furthermore, for cinerarin derived from cineraria (*Senecio cruentus*) which has one aliphatic acyl group and three aromatic acyl groups, it has been reported that the stability of the pigment decreases in a neutral aqueous solution by removing aromatic acyl groups (Goto et al., Tetrahedron 25: 6021, 1984). For gentiodelphin derived from gentians (*Gentiana makinoi*) also, it has been reported that an intra-molecular stacking of the sandwich type occurs due to the presence of two aromatic acyl groups in the molecule, which results in stabilization of the pigment in an aqueous solution (Yoshida et al., Tetrahedron 48: 4313, 1992). Moreover, Yoshida et al. have demonstrated that each of glucose at position 5 and glucose at position 3' of anthocyanin has an acyl group bound thereto (Tetrahedron 48: 4313, 1992). It has also been reported that anthocyanin in the leaves of perillas (*Perilla ocimoides*) is shisonin in which coumaric acid is bound to glucose at position 3 of cyanidin 3,5-diglucoside (Tetrahedron Letters 27: 2413-2416, 1978).

However, these studies have been carried out from the aspect of organic chemistry such as structural studies of natural pigments and not from the aspect of biochemistry such as efforts to isolate enzymes which transfer acyl groups.

Of the transferases which transfer acyl groups to anthocyanin pigments, there are many reports on the malonyl group transferases which transfer an aliphatic acyl, including those from a cell culture of parsley (Matern et al., Arch. Biochem. Biophys. 208: 233, 1981; Matern et al., Arch. Biochem. Biophys. 226: 206, 1983; Matern et al., Eur. J. Biochem. 133: 439, 1983), seedlings of *Cicer arietinum* (Koster et al., Arch. Biochem. Biophys. 234: 513, 1984), and the like.

Aromatic acyl transfer reaction was first reported for Silene, a member of Caryophyllaceae (Kamsteeg et al., Biochem. Physiol. Pflanzen 175: 403, 1980), and the activity of aromatic acyltransferase has similarly been found in the soluble enzyme fraction of *Matthiola* (Teusch et al., Phytochemistry 26: 991, 1986).

However, these reports have been limited to a mere demonstration of the presence of enzymatic activity, and neither the corresponding enzyme proteins have been specified nor findings have been obtained on the primary structure of the enzymes much less the genes encoding them. For other aromatic acyl transferases as well no reports have elucidated the primary structure of proteins or genes. Furthermore, there are no reports of examples in which the acylating reactions of anthocyanin pigments were positively used to expand the range of colors of flowers and to grow them, or examples in which acylation was used in an attempt to stabilize anthocyanins.

On the other hand, the biochemical pathway of synthesis of anthocyanins of *Petunia hybrida* has been well studied (Wiering, H. and de Vlaming, P. Inheritance and biochemistry of pigments. *Petunia*, P49-65 (1984), Griesbach, R.J., asen, S. and Leonhardt, B.A., *Phytochemistry*, 30: 1729-1731, 1991), and the presence of anthocyanins which contain an acyl group is known. As the acyl group of anthocyanins of *Petunia*, coumaric acid or caffeic acid is known. One molecule of coumaric acid or caffeic acid is bound to rutinoside at position 3 of anthocyanin, whose chemical structure, when the anthocyanidin is malvidin, has been assigned to 3-O-(6-O-(4-O-coumaroyl)- $\alpha$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin and 3-O-(6-O-(4-O-caffeoyl)- $\alpha$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin, respectively. However, there were no reports on anthocyanins having two acyl groups.

#### Disclosure of the Invention

The present invention relates to genes encoding proteins having aromatic acyl group transfer activity and the use thereof. Thus, with regard to said use, there is described a method for controlling an acyl group transfer reaction to flavonoids, preferably anthocyanins, which provides a possibility of developing a wide range of flower colors for a single species. In particular, said method is considered to be useful for imparting bluish tints to the existing color of flowers, because the absorption maximum of anthocyanin shifts to the long wavelength direction by transfer of aromatic acyl groups.

In order to realize the above technology, it is necessary to elucidate the identity of enzymes responsible for aromatic acyl transfer reactions and to separate the cDNA which encodes said enzymes. Furthermore, by utilizing the homology of genes it is possible to separate the genes of other acyl group transfer enzymes. Moreover, production of stable anthocyanin pigments can be realized by acylation since acylation leads to increased stability of anthocyanins.

The inventors have isolated an acyltransferase from petals of gentians and determined the primary structure thereof. Furthermore, using recombinant technology we have also isolated cDNA's of acyltransferases of gentians, petunia, perillas, and cinerarias, and determined the nucleotide sequences of the structural genes. Thus, the present invention provides DNA sequences encoding acyltransferases which are present in the petals of gentians, petunias, and cineraria, and leaves of perillas. Furthermore, the enzymes of the present invention can be used to change the colors of flowers by acylating the anthocyanin pigments and to increase stability of anthocyanins.

#### Specific Description

Genes encoding acyltransferases may be obtained, for example, as follows. Thus, first an acyltransferase is purified from petals of gentians. Prior to the present invention all attempts to purify aromatic acyltransferases have failed. The inventors of the present invention have succeeded in purifying said enzyme for the first time by employing various chromatographic methods, especially affinity chromatography using a resin (for example, Blue Sepharose (TM) resin, etc.) on which is immobilized, for example, Cibacron Blue 3GA.

Then, the partial amino acid sequence of the acyltransferase is elucidated using the conventional method and a synthetic nucleotide corresponding said amino acid sequence is prepared.

On the other hand, poly A + RNA is extracted from petals of the same gentian, from which double stranded cDNA is synthesized using the conventional method and a cDNA library is further produced. Using the above double stranded cDNA as the template a DNA fragment specific to the gene of acyltransferase is obtained by the PCR method using the synthetic DNA primers which were used for synthesis of said synthetic DNA and cDNA. Then, using this DNA fragment as a probe, the above mentioned cDNA library is screened to obtain positive clones. Plasmid DNA which is recovered from the clones are separated and their nucleotide sequences are determined. Then the amino acid sequence obtained from analysis of the purified acyltransferase and the amino acid sequence of the acyltransferase deduced from the DNA nucleotide sequence are compared to confirm that the above positive clone is the desired cDNA clone.

The inventors have also found petunia mutant (VM) a mutant strain of petunia var. *Surtinia* purple (VM) (Suntory Ltd.), in which the color of the flower has been changed from red purple to purple, and determined the structure of anthocyanins according to the method as described by, for example, Yoshida et al. (Yoshida et al., *Tetrahedron* 48: 4313, 1992).

As the DNA of the present invention there is mentioned DNA encoding the amino acid sequence as set forth in any of SEQ ID No. 1 to 6. However, it is known that proteins having modified amino acid sequences in which several amino acids have been added, removed and/or replaced with other amino acids have enzymatic activity similar to the original protein. Accordingly, genes encoding proteins which have modified amino acid sequences wherein one or more amino

acids have been added, removed and/or replaced with other amino acids and which retain aromatic acyl group transfer activity are encompassed in the present invention.

The present invention also relates to genes encoding proteins which hybridize with the nucleotide sequence as set forth in any of SEQ ID No. 1 to 6 or a portion thereof, for example the portion encoding six or more amino acids of the consensus region, under the condition of, for example 2 to 5 x SSC and 50°C, and which have acyl group transfer activity. Furthermore, the optimum hybridization temperature depends on the nucleotide sequence and its length. Preferably the hybridization temperature becomes low, as the nucleotide sequence becomes short. For example, in the case of the nucleotide sequence (18 bases) encoding six amino acids, a temperature of 50°C or lower is preferred. The present invention also relates to genes encoding proteins having the amino acid sequence which has a homology of 15% or higher, preferably 25% or higher, for example 30% or higher with the amino acid sequence as set forth in any of SEQ ID No. 1 to 6, and which has aromatic acyl group transfer activity.

The DNA which has the original nucleotide sequence is obtained, as specifically described in Examples, by screening, for example a cDNA library.

DNA encoding the enzyme having a modified amino acid sequence can be synthesized by the conventional site-directed mutagenesis or a PCR method based on the DNA having the original nucleotide sequence. For example, a DNA fragment having a site which is desired to be modified is obtained by digestion with restriction enzymes of cDNA or genomic DNA obtained as above, which is then used as the template to obtain the DNA fragment having the desired modification inserted therein by site-directed mutagenesis or a PCR method, and by ligating this to the DNA which encodes other parts of the desired enzyme.

Alternatively, in order to obtain DNA encoding an enzyme having a shortened amino acid sequence, DNA encoding an amino acid sequence longer than the desired amino acid sequence, for example DNA encoding the full-length amino acid sequence is cut with the desired restriction enzyme. When the resulting DNA fragment does not encode the desired entire amino acid sequence, the missing portion can be complemented by ligating synthetic DNA.

A gene encoding acyltransferase according to the present invention can be obtained by expressing the above clone in *Escherichia coli* and yeast using gene expression systems, confirming that the gene obtained encodes acyltransferase, and elucidating the translation region of the gene of acyltransferase. Furthermore, by expressing said gene a genetic product, the protein of the desired acyltransferase can be obtained.

Alternatively, it is also possible to obtain said protein using an antibody against the amino acid sequence described in any of SEQ ID No. 1 to 6.

Thus, the present invention relates to a recombinant vector comprising said DNA, in particular an expression vector, and a host transformed with said vector. As the host, a eukaryotic or prokaryotic organism may be employed. The prokaryotic organisms which may be used include a bacterium belonging to the genus *Escherichia*, for example, *Escherichia coli* a bacterium belonging to the genus *Bacillus*, for example *Bacillus subtilis*, or any other conventional hosts.

The eukaryotic organisms which may be used include lower eukaryotes, for example eukaryotic microorganisms, for example fungi such as yeast or filamentous fungi. As the yeast, there are mentioned *Saccharomyces* such as *Saccharomyces cerevisiae*, and as the filamentous microorganisms, there are mentioned *Aspergillus* such as *Aspergillus oryzae* and *Aspergillus niger*, and *Penicillium*, and the like. Moreover, animal cells or plant cells may be used. The animal cells which may be used include cell lines of mouse, hamster, monkey, human, and the like. Furthermore, insect cells such as silkworm cells or larvae of silkworm themselves may be used as a host.

The expression vectors of the present invention contain expression regulating regions, for example promoter and terminator, replication origin, and the like depending on the kind of the host to which they are introduced. As a promoter for bacterial expression vectors, conventionally used promoters such as *trc* promoter, *tac* promoter, *lac* promoter, etc. may be used. As a promoter for yeast, for example, glyceraldehyde-3-phosphate dehydrogenase promoter, PH05 promoter, and the like may be used. As a promoter for filamentous organisms, for example, *anylase*, *trp C*, and the like may be used. As a promoter for animal cell hosts, viral promoters such as SV40 early promoter, SV40 late promoter, and the like may be used.

The construction of an expression vector may be carried out according to a conventional method using restriction enzymes, ligase, and the like. Transformation of hosts with an expression vector may also be carried out according to a conventional method.

In the manufacture of said proteins, the desired protein can be obtained by culturing, growing, or breeding a host transformed with the above-mentioned expression vector, and then subjecting the culture to gel-filtration, centrifugation, cell disruption, gel-filtration chromatography, ion exchange chromatography, and the like to recover and/or purify said protein.

Although the invention has been described with specific reference to acyltransferases derived from gentians, petunias, perillias, and cinerarias, it should be noted that the purification method of said enzyme may be wholly or partially modified to purify acyltransferases of other plants and then the amino acid sequences of said enzymes are determined in order to clone genes encoding said enzymes. By using as a probe cDNA of acyltransferase derived from a gentian according to the present invention, it was also possible to obtain cDNA of another acyltransferase from a gentian and

cDNA of another acyltransferase from a petunia. Accordingly, by using part or all of the gene of acyltransferase it is possible to obtain the gene of another acyltransferase. Comparison of these amino acid sequences revealed the presence of a region of a conserved amino acid sequence. By using this region it was also possible to obtain cDNA of acyltransferase of a perilla and a cineraria. A similar method can be applied to other plants to obtain cDNA or chromosomal DNA clone of a similar acyltransferase.

As has been described hereinabove, by purifying acyltransferases derived from a gentian, a petunia, a perilla and a cineraria and then obtaining antibody against said enzyme according to a conventional method, it is possible to clone cDNA or chromosomal DNA which produces a protein capable of reacting with said antibody. Thus, the present invention is not limited to acyltransferases derived from gentians, petunias, perillas and cinerarias, but relates broadly to aromatic acyltransferases.

Furthermore, the present invention relates to plants of which colors have been controlled by introducing gene of acyltransferase therein, or progenies thereof or their tissues, and they may be in the form of cut flowers.

Furthermore, in the present specification, CoA esters such as p-coumaroyl-CoA or caffeoyl-CoA etc. were mentioned as a donor of an acyl group in the acyl group transfer reaction of flavonoids involving anthocyanins, further more p-coumaroyl, feruloyl, or hydroxycinnamoyl-1-O-glucose such as sinapoyl-1-O-glucose can also be used as a donor of an aromatic acyl group (Glasgesen and Seitz, *Planta* 186: 582, 1992), and therefore enzymes according to the present invention can be used.

#### Examples

The present invention is now explained with reference to the following specific embodiments. The experimental procedures used were according to Molecular Cloning by Sambrook (Cold Spring Harbor Laboratory Press, 1989), unless otherwise specified.

#### Example 1. Search of acyltransferase from plants

##### (1) Preparation of substrate

Delphinidin 3,5-diglucoside and cyanidin 3,5-diglucoside were obtained from petals of Tapan violet (Suntory Ltd.), a breed of *Verbena hybrida*, by extracting a diacetylated form of each of the above and then by deacetylating them. Petals (348 g) of Tapan violet were homogenized with liquid nitrogen in a homogenizer, immersed in 1.5 L of 50% (v/v) acetonitrile and 0.2% trifluoro acetic acid (TFA), and then allowed to stand for three days.

The product thus obtained was filtered under aspiration through diatomaceous earth (#100) spread over filter paper then concentrated to half the volume in a rotary evaporator, followed by gel-filtration with HP-20 (Pharmacia). After washing with 800 ml of distilled water, the pigment fraction was eluted with 800 ml of 50% acetonitrile and 0.1% TFA. After concentration in an evaporator, it was lyophilized to obtain crude pigment (7.3 g).

Since the main pigments in Tapan are 3,5-diacetylglucoside of delphinidin and cyanidin, the following procedure of deacetylation was carried out. One gram of the crude pigment was dissolved in 50 ml of methanol and aerated with nitrogen gas for 15 minutes to expel dissolved oxygen and then cooled on ice.

Separately, dissolved oxygen was similarly expelled from 50 ml of 1 N sodium hydroxide, into which the above pigment solution was added dropwise while stirring in the ice, and was stirred for further 30 minutes to effect hydrolysis. One ml of 6 N hydrochloric acid was added thereto to stop the reaction. Then, 5 ml of distilled water was added and concentrated to half the volume in an evaporator, to which methanol was added to a final concentration of 10%. Two ml aliquots were applied to Sep Pac C18 column (Waters Association), which was then washed with 5 ml of distilled water, and eluted with 2 ml of 30% acetonitrile and 0.6% TFA.

All eluates were collected and concentrated in an evaporator, and then fractionated by HPLC. Using a DEVELOSIL ODS-10/20 (50 x 300 mm; Nomura Kagaku K.K.) column, elution was effected at a linear gradient of TFA from 0.1% to 0.3% and acetonitrile from 10% to 30% over 120 minutes. Fractions were collected every 0.5 minute at a flow rate of 32 ml per minute. Absorption spectrum of each pigment fraction was measured to separate delphinidin-3,5-diglucoside and cyanidin 3,5-diglucoside, which were then concentrated and lyophilized (delphinidin-3,5-diglucoside, 75 mg and cyanidin 3,5-diglucoside, 50 mg). They were each dissolved in 0.5% TFA to a concentration of 1.5 mg/ml and stored at -80°C until use.

Synthesis of another substrate, hydroxy cinnamoyl-CoA was carried out in the following manner. First, an ester was synthesized from caffeic acid (Nakalai tesque) and N-hydroxysuccinimide (Merck) according to a literature (Stockigt and Zenk, *Z. Naturforsch.* 30: 352, 1975). This ester (0.5 mmol) was dissolved in 2 ml of acetone. Separately 0.1 mmol of Coenzyme A (CoA: KOHJIN) and 1 mmol of sodium hydrogen carbonate were dissolved in 20 ml of water, to which was added dropwise the ester solution prepared above.

After the mixture was reacted overnight while stirring under nitrogen gas at room temperature, it was concentrated in a rotary evaporator and centrifuged (27,000 x g, 10 min) to remove insoluble matter and the desired product was col-

lected using HPLC. Using a DEVELOSIL ODS-10/20 (50 x 300 mm; Nomura Kagaku K.K.) column, elution was carried out at a linear gradient of acetonitrile from 18% to 36% in the presence of 0.1% of TFA over 40 minutes. Fractions were collected every 0.8 minute at a flow rate of 32 ml per minute. The absorption spectrum of each fraction was measured (200 to 400 nm) to collect the fractions having an absorption maximum in the range of 344 to 348 nm as the caffeoyl CoA fraction. After concentration in a rotary evaporator, they were separated using the same column again.

However, elution was carried out by isocratic chromatography of 18% acetonitrile and 0.1% TFA, and the absorption spectrum was measured simultaneously to concentrate the fractions containing the desired compounds in a rotary evaporator, which were then lyophilized. This method produced 35  $\mu$ mol of the products. By substituting coumaric acid for caffeic acid above, p-coumaroyl-CoA was synthesized. The product was dissolved in distilled water at 2 mg/ml and stored at -80°C until use.

## (2) Extraction method of the crude enzyme solution

Three grams of the plant tissue (petals, edible parts, etc.) from which enzyme was to be extracted was frozen in liquid nitrogen and was homogenized in a mortar. It was further homogenized by adding 10 ml of the extraction buffer (100 mM phosphate buffer, pH 7.5, 10 mM sodium ascorbate, 14 mM 2-mercaptoethanol) and was filtered through three layers of gauze. After adding 3 g of DOWEX (1-X2, 100-200 mesh; Muromachi Kagaku Kogyo K.K.) and stirred for 10 minutes, the resin was removed by filtration under aspiration and the debris of the plant tissue was removed by centrifugation (27,000 x g, 20 minutes). It was then subjected to salting out under 70% saturated ammonium sulfate to precipitate proteins. The precipitate was suspended into 1 ml of the solubilizing buffer (20 mM phosphate buffer, pH 7.5, 14 mM 2-mercaptoethanol) and insoluble matter was removed by centrifugation (27,000 x g, 5 minutes). Then it was desalted using Sephadex G-25 column (NAP-10; Pharmacia) which had been equilibrated with the solubilizing buffer and the solution thus obtained was used as the crude enzyme solution.

## (3) Method of measuring enzyme activity

Fifty  $\mu$ l of a reaction mixture containing 100 mM phosphate buffer, pH 8.5, 24 nmol of delphinidin 3,5-diglucoside, 21.5 nmol of caffeoyl-CoA, and 20  $\mu$ l of the enzyme solution was reacted at 30°C for 10 minutes. After stopping the reaction by adding 50  $\mu$ l of acetonitrile containing 13.8% (v/v) acetic acid and insoluble matter was removed by centrifugation (18,000 x g, 5 minutes), it was analyzed by high performance liquid chromatography (HPLC). Twenty  $\mu$ l of the reaction mixture was analyzed using a C18 reverse phase column (YMC-Pack ODS-A, 6.0 x 150 mm; YMC) and 21.6% acetonitrile and 0.1% trifluoroacetic acid at a flow rate of 1 ml per minute. The compounds were detected using a three dimensional chromatography system (CLASS-LC10; Shimadzu Seisakusho, K.K.) and it was found that the product has an absorption maximum at about 330 nm which is absent in the substrate and that the absorption maximum in the visible light range shifted by about 6 nm from 519 nm to 525 nm, confirming that an acyl group (caffeic acid) is bound, and delphinidin 3-glucosyl 5-caffeoyl glucoside has been produced.

By detecting at a wavelength of 520 nm, the ratio of the peak area of the product (delphinidin 3-glucosyl 5-caffeoyl glucoside) to the sum of the peak areas of the substrate (delphinidin 3,5-diglucoside) and the product (delphinidin 3-glucosyl 5-caffeoyl glucoside) was determined to calculate the mole number of the product, which was defined as the enzymatic activity (kat). The retention time for each compound in this HPLC analysis was as follows: caffeoyl-CoA, 6.3 min; delphinidin 3,5-diglucoside, 3.3 min; delphinidin 3-glucosyl 5-caffeoylglucoside, 5.3 min.

Since under this reaction condition delphinidin 3,5-diglucoside in the reaction mixture is modified with caffeic acid by the action of acyltransferase resulting in color change of the reaction mixture from dark blue to reddish purple, the enzymatic activity can be determined, as a simple method, by carrying out the reaction in a microtiter plate.

When the plate after the reaction is allowed to stand at room temperature for a prolonged period of time (one day to one week), delphinidin 3,5-diglucoside which was not acylated becomes colorless, whereas the delphinidin 3,5-diglucoside which was acylated by the action of the enzyme retains the reddish purple color, so that stabilization of delphinidin 3,5-diglucoside in a neutral to alkaline solution was observed because of its acylation. Similarly, when cyanidine 3,5-diglucoside was used as the substrate, the color of the reaction mixture changed from reddish purple to dark blue and the pigment becomes stabilized, and hence it is possible to detect the enzymatic activity by a simple enzyme assay.

On the other hand, when caffeoyl-CoA is replaced with p-coumaroyl-CoA, acylation-derived color change and stabilization of anthocyanin are observed, but the degree of change in the tone of color is smaller than with caffeoyl-CoA.

## (4) Search for acyltransferase

Crude enzyme solutions were extracted from a variety of plants including gentians, iris, delphiniums, stocks, *Eustoma russellianum* Griseb., pinks, sweet peas, Larkspurs, pansy, cinerarias (petals for the above plants), red cabbages, red onions, Kintoki carrots, western carrots, purple potatoes, perillas (edible parts for the above plants) and egg plants (epithelial part of the fruit), and their enzymatic activities were determined. As a result, acyltransferase activities of 0.63,

0.0012, and 21.8 nkat/mg protein were detected in the extracts from *Eustoma russellianum* Griseb. pinks, and gentians, respectively. Gentian, which had the highest acyltransferase activity per protein extracted, was used as a material for enzyme purification.

Determination of protein concentration was carried out using the Bio-Rad Protein Assay (Bio-Rad).

## Example 2. Purification of acyltransferase derived from gentians

### (1) Purification of enzyme

Enzyme was extracted from petals of *Gentiana triflora* var. japonica. The following experiment was carried out at 0 to 4°C unless otherwise noted. Three kilograms of petals of *Gentiana triflora* var. japonica was homogenized in the presence of liquid nitrogen using the Excell Auto Homogenizer (DX-3; Nihon Seiki Seisakusho). After adding 8 L of the extraction buffer (100 mM phosphate buffer, pH 7.0, 10 mM sodium ascorbate, 10 µM p-aminodiphenyl methanesulfonyl fluoride hydrochloride (p-APMSF; Wako Pure Chemicals K.K.)), 5 mM dithiothreitol (DTT; Nakalai Tesque), and 500 g of polycar SB-100 (Wako Pure Chemicals K.K.), it was completely pulverized.

After the pulverized liquid was squeezed with 4 layers of gauze, it was further centrifuged (11,000 x g, 30 min) to remove cell debris. Then it was salted out with 40% saturated ammonium sulfate and insoluble matter was removed prior to salting out again with 70% saturated ammonium sulfate. The precipitate was suspended into 250 ml of the solubilizing buffer (20 mM Tris-HCl, pH 7.0, 10 µM p-APMSF, 1 mM DTT), and insoluble matter was removed by centrifugation. Then it was desalted using Sephadex G-25 column (95 x 110 mm; Pharmacia) which had been equilibrated with the same buffer. The protein-containing fractions were collected (860 ml) and subjected to the following chromatography.

Each of chromatographies of Q-Sepharose Fast Flow, HiTrap Blue and Phenyl Superose were carried out using the FPLC system (Pharmacia).

First, the samples were applied to Q-Sepharose Fast Flow (26 x 100 mm; Pharmacia) which had been equilibrated with the same buffer. After adequately washing the column with the same buffer, it was eluted with a linear gradient of sodium chloride from 0 M to 0.4 M in 60 minutes (8 ml/min). After the fractions containing enzymatic activity were pooled (130 ml), they were subjected to affinity chromatography. It was then applied to three columns of HiTrap Blue (5 ml, 16 x 25 mm; Pharmacia) connected in a series, adequately washed with the same buffer, and eluted with the same buffer containing 1 M sodium chloride. The active fractions were salted out with 70% saturated ammonium sulfate to obtain a protein precipitate.

The precipitate was suspended in 1 ml of the solubilizing buffer and insoluble matter was removed by centrifugation, and then was applied to Sephacryl S-200 (25 x 1150 mm; Pharmacia) which had been equilibrated with the solubilizing buffer. At a flow rate of 0.2 ml per minute about 3 ml fractions were collected and after the active fractions were collected again (27 ml), sodium ammonium was added thereto to a concentration of 1 M. After fully stirring, it was centrifuged (39,000 x g, 10 min) to remove insoluble matter and applied to Phenyl Superose 5/5 (5.0 x 50 mm; Pharmacia) which had been equilibrated with the solubilizing buffer containing 1 M sodium ammonium.

After adequately washing at a flow rate of 0.5 ml, the concentration of sodium ammonium was linearly decreased from 1 M to 0 M over 60 minutes to elute protein. A 0.5 ml aliquot of each fraction was measured for enzymatic activity. In an analysis by SDS-polyacrylamide gel electrophoresis a band of molecular weight about 50,000 was observed as an almost single protein and since correlation was observed between this protein and activity, the protein was concluded to be the desired acyltransferase. The fractions (12 ml) having activity were further purified by reverse phase HPLC in order to obtain a single product.

Using a DEVELOSIL 300 C4-HG-5 (4.6 x 250 mm; Nomura Kagaku K.K.) column, elution was carried out at a linear gradient of acetonitrile from 40.5% to 56.7% in the presence of 0.1% trifluoroacetic acid over 30 minutes at a flow rate of 1 ml per minute. One ml fractions were collected while monitoring absorbance at 280 nm. Each fraction was further analyzed by SDS-polyacrylamide gel electrophoresis to collect fractions containing protein of molecular weight about 50,000. By repeating this HPLC for 30 times and concentrating in a speed Vac Concentrator (Savant), about 0.2 ml of single protein product was obtained.

### (2) Analysis of purified protein

When 500 pmol of purified product was subjected to the amino acid sequencer (PSQ-1; Shimadzu Seisakusho K.K.), 200 pmol of glutamic acid at the first stage of Edman degradation and 90 pmol of glutamic acid at the second stage were detected, but not at the third stage and thereafter. Accordingly, it was inferred that the N-terminal of the enzyme was blocked in some way or other.

However, since it is known that when the N-terminal is glutamic acid, pyroglutamic acid is formed and the sequence as described above by Edman degradation is observed, it is highly probable that the N-terminal of the enzyme is glutamic acid.

The remainder of the precipitate was dissolved in a solution containing 80  $\mu$ l of 45 mM Tris-HCl, pH 8.5, 3.6 M urea, and 0.09% SDS, to which was added 16 pmol of lysyl endopeptidase (Lysyl Endopeptidase: derived from Achromobacter lyticus; Wako Pure Chemicals K.K.) and was reacted at 37°C for 6 hours. The reaction mixture was separated directly by a DEVELOSIL 300 C4-HG-5 column.

The separation condition was a flow rate of 0.7 ml of a linear gradient of acetonitrile from 0% to 80% over 70 minutes in the presence of 0.1% trifluoroacetic acid. By monitoring absorbance at 210 nm fragments having a peak absorbance were collected. Out of 13 peak fractions thus obtained, three fractions which eluted at acetonitrile concentrations of 32% to 40% were concentrated in the Speed Vac Concentrator and then separated and purified using an ODS column (DEVELOSIL 300 ODS-HG-5; Nomura Kagaku K.K.) under the same condition as above.

Each fraction was concentrated to dryness in the Speed Vac Concentrator, dissolved in 30  $\mu$ l of 40% acetonitrile, and then subjected to the amino acid sequencer. As a result, the amino acid sequences of six peptides were able to be analyzed. The amino acid sequence for each peptide is shown below (the sequence is shown in the direction from the amino terminus to the carboxy terminus):

Amino acid sequence (AT73): Arg-Phe-Leu-Gly-Ile-Thr-Gly-Ser-Pro-Lys (SEQ ID No. 7)

Amino acid sequence (AT72): Ile-His-Met-Asp-Ala-Phe-Ala-Lys (SEQ ID No. 8)

Amino acid sequence (AT741-1): Gly-Val-Glu-Ile-Gly-Val-Ser-Leu-Pro-Lys (SEQ ID No. 9)

Amino acid sequence (AT741-2): Ala-Ser-Leu-Ser-Leu-Thr-Leu-Lys (SEQ ID No. 10)

Amino acid sequence (AT9): His-Tyr-Val-Pro-Leu-Ser-Gly-Asn-Leu-Leu-Met-Pro-Ile-Lys (SEQ ID No. 11)

Amino acid sequence (AT83): Val-Arg-Ala-Thr-Tyr-Val-Leu-Ser-Leu-Ala-Glu-Ile-Gln-Lys (SEQ ID No. 12)

### Example 3. cDNA cloning of acyltransferase derived from gentians (1)

#### (1) Construction of cDNA library

Petals were collected from commercial gentians (*Gentiana triflora* var. *japonica*) and homogenized under liquid nitrogen in a mortar. From the homogenate, RNA was obtained by the method utilizing guanidine thiocyanate/cecum chloride and then poly A + RNA was obtained using the Oligotex (Nihon Roche) in the method recommended by the manufacturer. The method using guanidine thiocyanate/cecum chloride was carried out according to the method described in detail in R. McGookin, Robert J. Slater et al., Methods in Molecular Biology vol. 2 (Human Press Inc. 1984).

Using the obtained poly A + RNA as the template, double stranded cDNA was synthesized using the ZAP-cDNA synthesis kit (manufactured by Stratagene) and was cloned into phage vector  $\lambda$ ZAP11. Furthermore, using the Gigapack11 Gold Packaging Extract kit of the same company cDNA library was constructed by the method described in the kit insert.

#### (2) Designing synthetic DNA primers

Among the amino acid sequences obtained in Example 2, the sequence represented by Ile-His-Met-Asp-Ala-Phe-Ala-Lys (SEQ ID No. 13) is very likely to be Lys-Ile-His-Met-Asp-Ala-Phe-Ala-Lys (SEQ ID No. 14) considering the specificity of lysyl endopeptidase. Using the portion represented by the amino acid sequence: Lys-Ile-His-Met-Asp-Ala-Phe-Ala (SEQ ID No. 15) in this sequence, the following oligonucleotide was synthesized:

Nucleotide sequence (Oligo 1): 5'-AARATHCAYATGGAYGCITTYGC-3' (SEQ ID No. 16)

Here, the sequence of nucleic acids is shown by the one-letter code in accordance with IUPAC-IBU. That is, A: adenine, C: cytosine, G: guanine, T: thymine, Y: C or T, R: A or G, H: A or C or T, and I: inosine.

Furthermore, another oligonucleotide shown below was also synthesized based on the primer used for construction of the cDNA library mentioned above:

Nucleotide sequence (Oligo 2): 5'-CTCGAGTTTTTTTTTTTTTTT-3' (SEQ ID No. 17)

#### (3) Cloning of fragments of the acyltransferase gene

Using about 0.1  $\mu$ g of double stranded cDNA derived from RNA of petals of gentians and Oligo 1 and Oligo 2 as primers, the PCR reaction was carried out. The reaction was carried out using the polymerase chain reaction kit Gene Amp (Takara Shuzo K.K.) for 35 cycles with one cycle comprising 95°C for 1 minute, 45°C for 1 minute, and 72°C for 2 minutes. When the reaction product thus obtained was run on a 1% agarose gel electrophoresis, a specific DNA fragment of about 400 bp was observed. This DNA fragment was recovered and 10 ng thereof was subjected to 25 cycles of the above-mentioned polymerase chain reaction using the DIG-nucleotide mixture (Boehringer) and synthetic nucleotide I and II to obtain DIG-labelled DNA fragments.



## (4) Cloning of cDNA of acyltransferase

$\lambda$  phage library obtained as above was infected to *E. coli* strain XL1-Blue (Stratagene) to screen five plates (diameter, 13.5 cm) containing 50,000 plaques per plate.

Phage was adsorbed to a filter (Hybond N+, Amersham) and treated in the method recommended by the manufacturer, and then the filter was allowed to remain in the hybridization buffer (5 x SSC, 50% formamide, 50 mM sodium phosphate buffer, pH 7.0, 7% SDS, 2% Blocking reagent (Boehringer), 0.1% lauryl sarcosine, 80 mg/ml salmon sperm DNA) at 42°C for 1 hour. The DIG-labelled DNA fragment obtained above was added to the hybridization solution and incubated for 16 hours.

The filter was washed with a washing solution (0.2 x SSC, 0.1% SDS) and then an enzyme immunoassay (Boehringer Mannheim) using the DIG-specific antibody labelled with alkaline phosphatase was carried out to detect by color development using 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium salt. The detection method used was as set forth in the manufacturer's instructions.

As a result, a few dozen positive clones were obtained. From 20 clones of them cDNA was collected on the plasmid pBluescript SK-. Insertion of cDNA was examined by agarose gel electrophoresis and it was found that the insertion of cDNA's of varying sizes was observed in all clones and the longest among them was 1.7 kb. Among them 9 clones were chosen and were subjected to analysis by restriction enzymes. Accordingly, it was found that similar patterns of restriction enzyme cleavage were observed, though their sizes were varied.

## (5) Determination of nucleotide sequence

Plasmid was extracted from the clones thus obtained. Using the AB1373A DNA Sequencer (Perkin Elmer), for six clones (pGAT2, pGAT3, pGAT4, pGAT7, pGAT8, and pGAT11) out of nine which are considered to contain the full-length, the nucleotide sequence of the 5' end of cDNA was determined by the dideoxy sequence method using the fluorogenic reagents recommended by the same manufacturer.

The result suggested that these clones have the same nucleotide sequence and differ in the length of cDNA. From among these clones, the entire nucleotide sequence of pGAT4 was determined. Determination of the nucleotide sequence was carried out for each clone after a series of deleted clones were obtained using the Deletion Kit for Kilo-Sequence (Takara Shuzo, K.K.).

## (6) Comparison of the nucleotide sequence with the amino acid sequence

cDNA which was inserted into pGAT4 represented 1703 bases, which was found to contain an open reading frame (ORF) comprising 1410 bases (containing the stop codon). The sequence is shown in the sequence listing SEQ ID No.

1. Since all of the partial amino acid sequences of the acyltransferase revealed in Example 2 occurred as amino acid sequences in the ORF, it was concluded that the cloned cDNA was the gene of acyltransferase derived from gentians. Analysis of the amino terminal of the initiation codon suggested that glutamic acid is the residue of the amino terminal, so that it was inferred that the first ATG from the 5' end was the initiation codon on the nucleotide sequence of the cDNA.

On the other hand, since the cDNA of pGAT8 is shorter than pGAT4 by 7 bases at the 5' end, it was suggested that this was not the full-length cDNA.

Example 4. Expression of genes in *E. coli*

## (1) Construction of expression plasmid

pTrc99A (Pharmacia), an *E. coli* expression vector, was used for expression of the acyltransferase gene of *E. coli*. This pTrc99A contains *E. coli* trc promoter which can be induced by isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) and therefore by inserting the desired gene downstream of said promoter the gene can be expressed in *E. coli*.

A restriction enzyme NcoI site has been inserted thereto by making use of the initiation codon, ATG sequence, so that direct expression of the desired gene from the initiation codon is possible by recombining it with NcoI.

pGAT10 was constructed by recombining the 1.8 kb DNA fragment (containing all nucleotide sequences as set forth in SEQ ID No. 1) obtained by digestion of pGAT4 with EcoRI and KpnI which are present in the present vector with the EcoRI and KpnI sites of the above-mentioned pTrc99A.

In order to introduce a NcoI site in the vicinity of the initiation codon of the acyltransferase, the following two oligonucleotides were synthesized which correspond to the vicinity of the initiation codon and the inside of the acyltransferase (about 300 bases from the initiation codon):

Oligonucleotide (GAT-NcoI): 5'-TTCACCATGGAGCAAATCCAAATGGT-3' (SEQ ID No. 18)

Oligonucleotide (GAT-ScaI): 5'-CGAGTCGCCCTCATCAC-3' (SEQ ID No. 19)

With 10 ng of pGAT4 as the template, a PCR reaction was carried out using the above oligonucleotides as the prim-

ers. The reaction was carried out using the polymerase chain reaction kit Gene Amp (Takara Shuzo K.K.) for 15 cycles with one cycle comprising 95°C for 1 minute, 56°C for 1 minute, and 72°C for 2 minutes. When the reaction product thus obtained was run on a 1% agarose gel electrophoresis, a specific DNA fragment of about 300 bp was observed. This DNA fragment was collected and cleaved with restriction enzymes NcoI and AatI. It was then ligated to a 6 kb fragment which was obtained by cleaving pGAT101 with NcoI and AatI to construct pGAT102. It was confirmed that the nucleotide sequence of the PCR-amplified portion was the same as that of pGAT4 after construction of pGAT102.

## (2) Expression of acyltransferase gene in *E. coli*

pGAT102 was used to transform *E. coli* MM294 (supE44 hsdR endA1 pro thi) (Messelson and Yuan, Nature, 217: 1110-, 1968). The host used here need not be specifically defined and may be any *E. coli* host which can be used as the host for transformation, and other strains (such as JM109, DH5, etc.) which are generally used for transformation and which are readily available to those skilled in the art can be employed. The method for transforming *E. coli* was as described by Hanahan (J. Mol. Biol., 166: 557-, 1983). The transformed *E. coli* was inoculated into 2 ml of LB medium (trypton 10 g, yeast extract 5 g, sodium chloride 10 g were dissolved in one liter of distilled water and pH was adjusted to 7.2 with sodium hydroxide) and incubated at 37°C overnight.

One ml of the culture liquid was inoculated into 10 ml of M9 medium (sodium hydrogen phosphate 0.6%, potassium dihydrogen phosphate 0.3%, sodium chloride 0.5%, ammonium chloride 0.1%, glucose 0.5%, magnesium sulfate 1 mM, vitamin B1 4 µg/ml, pH 7.2) to which were added 0.5% casamino acid and 50 µg/ml of ampicillin, and cultured at 37°C for 3 hours, and then 40 µl of 0.5 M IPTG (the final concentration, 2 mM) was added and culture was continued for 5 more hours. After harvesting the cells, they were washed with 30 mM Tris-HCl buffer, pH 7.5, containing 30 mM sodium chloride, and then the washed cells were suspended into 1 ml of the same buffer. To the cells were added 1 mg of lysozyme, 25 µl of 0.25 M EDTA, and allowed to stand at 0°C for 30 minutes. The cells were then frozen and thawed for three times to disrupt the cells.

After centrifugation at 15,000 rpm for 30 minutes, the supernatant obtained was used as a crude enzyme solution and the enzymatic activity thereof was determined in the method for determination of enzymatic activity as set forth in Example 1(3). In the microtiter plate method the acyl group transfer reaction was confirmed in *E. coli* to which pGAT102 was introduced. Accordingly, they were analyzed by HPLC. As a result it was found that in the *E. coli* into which pGAT102 was introduced, 18.3 nmol of delphinidin 3-glucosyl 5-cafeoyl glucoside was formed from 24 nmol of delphinidin 3,5-diglucoside and 21.5 nmol of cafeoyl-CoA.

Combining this result with the known fact that in anthocyanin of gentian the acyl group is bound to glucose at position 5 and position 3', it was revealed that the acyltransferase encoded by pGAT4 catalyzes the reaction of transferring an acyl group to glucose at position 5 of anthocyanin 3,5-diglucoside.

Furthermore, delphinidin 3,5-diglucoside which was acylated by acyltransferase produced in *E. coli* have also shown a stable color development when allowed to stand at room temperature for a prolonged period of time similarly to the one acylated by acyltransferase obtained by purifying from gentian.

## Example 5. Expression of genes in yeast

### (1) Expression vector in yeast

As the expression vector of yeast, pYE22m as described in Japanese Unexamined Patent Publication (Kokai) No. 4-228078 was used.

### (2) Expression of the acyltransferase gene in yeast

About 1.8 kb of DNA fragment obtained by digesting either pGAT4 or pGAT8 at restriction enzyme sites, EcoRI and KpnI, present in each of said vectors was ligated to about 8 kb of DNA fragment obtained by digesting similarly pYE22m at EcoRI and KpnI sites to construct yeast expression vectors pYGAT4 and pYGAT8. pYGAT4 starts translation at the first methionine, but pYGAT8 which lacks part of 5' end of the isolated cDNA starts translation not at the translation initiation methionine of acyltransferase (number of amino acid sequence in the sequence listing SEQ ID No: 1), but at the next methionine (number of amino acid sequence in the sequence listing SEQ ID No: 5).

In these yeast expression vectors, the cDNA encoding acyltransferase is ligated downstream of the promoter for glyceraldehyde-3 phosphate dehydrogenase, one of the constitutive yeast promoters, and its transcription has been regulated by said promoter.

Using the method by Ito et al. (Ito et al., J. Bacteriol., 153: 163-168, 1983), a yeast *Saccharomyces cerevisiae* G1315 (Ashihari et al., Appl. Microbiol. Biotechnol. 30, 515-520, 1989) was transformed. The transformed yeast was selected based on its recovery of synthetic ability of tryptophan.

It should be noted that the yeast host as used herein for transformation is not limited, but it may be any strain which

displays a tryptophan requirement due to its incomplete TRP1 gene (for example, one commercially available from the Yeast Genetic Stock Center; Berkely, CA, USA; Catalogue 7th edition (1991), page 36).

The transformant obtained was cultured under shaking in 10 ml of Burkholder medium (Burkholder, Amer. J. Bot. 30: 206-210) containing 1% casamino acid. As a control experiment, the yeast which has spontaneously recovered its synthetic ability of tryptophan was cultured in a similar manner.

After harvesting the cells, they were washed with the same amount of the cell disruption buffer (30 mM Tris-HCl, pH 7.5, 30 mM sodium chloride), and suspended further in 1 ml of the same buffer and then was transferred into a 1.5 ml Eppendorf tube. After centrifugation, the supernatant was removed and the precipitated cells were resuspended into 0.4 ml of the same buffer, to which was added 400 mg of glass beads (Glass Beads 425-600 microns Acid-Wash, Sigma) and shaken vigorously to disrupt the yeast cells.

The supernatant after centrifugation was used as a crude enzyme solution and the enzymatic activity thereof was determined by the method for determination of enzymatic activity as set forth in Example 1(3). Since acyl group transfer reaction was observed by the microtiter plate method in all yeasts into which pYGAT4 and pYGAT8 were introduced, they were then analyzed by HPLC. The yeast used as the control did not show any activity of acyl group transfer.

The result indicated that 16.6 nmol and 20.9 nmol of delphinidin 3-glucosyl 5-caffeoyl glucoside were formed from 24 nmol of delphinidin 3,5-diglucoside and 21.5 nmol of caffeoyl-CoA, respectively in the yeast into which pYGAT4 and pYGAT8 were introduced. Both of the proteins which were produced by pYGAT4 and pYGAT8 had acyl group transfer activity though their amino termini were different.

Furthermore, delphinidin 3,5-diglucoside which was acylated by the acyltransferase produced by the yeast have shown a stable coloration even when allowed to stand at room temperature for a prolonged period of time similarly to the one acylated by the acyltransferase obtained by purifying from gentians.

#### Example 6. cDNA cloning of the acyltransferase derived from gentian (2)

Among the DNA fragments obtained by digestion of pGAT4 as set forth in Example 3(6), i.e. pGAT4 having the DNA as set forth in the sequence listing SEQ ID No. 1, with restriction enzymes EcoRI and NdeI, two DNA fragments which contain the translation region of acyltransferase were collected together and labelled with DIG as in the methods mentioned above. Using this as a probe, the phage of cDNA library from petals of gentians was adsorbed onto a filter (Hybond N+, Amersham), which was then regenerated by removing the pigments and the DIG labels attached to the filter according to the method recommended by the manufacturer (Amersham) and subjected to hybridization in a low concentration formamide hybridization buffer (5 x SSC, 30% formamide, 50 mM Tris-HCl, pH 7.5, 1% SDS) at 42°C for 16 hours.

After washing at 50°C in the wash solution (5 x SSC, 0.1% SDS), the filter was allowed to develop color as described in Example 3(4). A few dozen clones developed color. Out of the clones which developed color, 12 clones which did not develop color in Example 3(4) were obtained. The nucleotide sequences of cDNA of these clones were determined from the 5' end in the above-mentioned method to find that the nucleotide sequences of 11 clones coincided with that of pGAT4, but one clone did not, which was designated as pGAT106.

The entire nucleotide sequence of pGAT106 was determined as described above. The cDNA introduced into pGAT106 represented 1622 bases in length, in which an ORF comprising 1440 bases (containing the stop codon) was found. This is shown in the sequence listing SEQ ID No. 2. For the ORF contained in SEQ ID No. 2, its homology was examined with the entire region of the amino acid sequence encoded by pGAT4. The homology was found to be 38%.

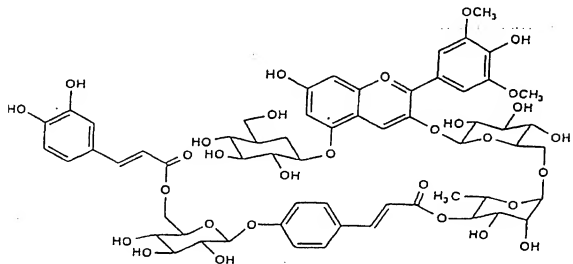
Since the amino acid sequence encoded by pGAT106 is homologous with that of the enzyme encoded by pGAT4, it is inferred that the former has a similar enzymatic activity, i.e. an activity of catalyzing acyl group transfer to anthocyanins. The fact that the anthocyanin of gentians has acyl groups at glucose molecules at position 5 and 3' suggests that pGAT106 catalyzes the enzymatic reaction of transferring an acyl group to position 3' of anthocyanin. The result indicates that acyltransferases may be different in the positions of sugars of anthocyanins which they transfer an acyl group but that the amino acid sequences and the nucleotide sequences encoding them are homologous.

As hereinabove described, there are many anthocyanins which have acyl groups, and these compounds vary widely in the number and position of acyl groups. Accordingly, it is expected that there are a number of enzymes which catalyze acyl group transfer reaction. It is readily inferred that the amino acid sequences of these enzymes have homology with the amino acid sequences of pGAT4 and pGAT106 obtained herein. Based on this other acyltransferase genes can be obtained.

#### Example 7. Anthocyanin of petunias

Anthocyanins found in a mutant (VM) in which the color of the flower was changed to purple from the original red-dish purple of *Petunia hybrida* var. Surfina purple (Suntory Ltd.), a breed of, was extracted by pulverizing petals thereof in liquid nitrogen and then extracting with 50% acetonitrile and 0.1% TFA. After filtration, the filtrate was separated and purified by ODS and ODP reverse phase column chromatographies. When one of the compounds was analyzed in

detail by FABMS,  $^1\text{H}$  NMR, and  $^{13}\text{C}$  NMR, a new anthocyanin was found. Its structure is shown below:



That is, the structure was

3-O-(6-O-(4-O-(4-O-(6-O-caffeoyl- $\beta$ -D-glucopyranosyl)-coumaroyl)- $\alpha$ -L-rhamnosyl)- $\beta$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin, or an anthocyanin to which two acyl groups are bound.

In addition,

3-O-(6-O-(4-O-(4-O-(6-O-coumaroyl- $\beta$ -D-glucopyranosyl)-coumaroyl)- $\alpha$ -L-rhamnosyl)- $\beta$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin,

3-O-(6-O-(4-O-(4-O-(6-O-caffeoyl- $\beta$ -D-glucopyranosyl)-caffeoyl)- $\alpha$ -L-rhamnosyl)- $\beta$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin,

3-O-(6-O-(4-O-(4-O-(6-O-coumaroyl- $\beta$ -D-glucopyranosyl)-cafeoyl)- $\alpha$ -L-rhamnosyl)- $\beta$ -D-glucopyranosyl)-5-O- $\beta$ -D-glucopyranosyl-malvidin were also detected. The anthocyanins were found to be present in the dark purple petals of Fulcon Blue (Sakata Seed Corp.), Old Glory Blue (Ball Seeds), and the like. Thus, anthocyanins having two acyl groups conceivably contribute to the dark purple color of petunias.

Accordingly, the foregoing suggests that acyltransferases related to anthocyanins derived from petunia come in two types: the enzyme which catalyzes a reaction of transferring coumaric acid or caffeic acid to rutinoid at position 3 of anthocyanin, and the enzyme which catalyzes a reaction of transferring coumaric acid or caffeic acid to monoacyl malvidin via glucose.

#### Example 8. cDNA cloning of the acyltransferase derived from petunias

cDNA portion of pGAT4 described in Example 3(6), i.e. pGAT4 having DNA as set forth in the sequence listing SEQ ID No. 1, was labelled with DIG in the method described above, and the cDNA library of petals of *Petunia hybrida* var. Old Glory Blue was screened by the plaque hybridization technique. Hybridization and washing were carried out under the condition similar to the one as set forth in Example 6.

About 200,000 clones were screened, from which one weakly hybridizing clone was obtained. This clone was designated as pPAT5. Determination of the nucleotide sequence revealed that more than one DNA were inserted in pPAT5. Thus, there was a sequence similar to that of C-terminal of the protein encoded by pGAT4 and pGAT106 in the reverse primer side of the plasmid. Based on the reverse primer, a nucleotide sequence: 5'-AACAGCTATGACCATG-3' (SEQ ID No. 20) was synthesized and the oligonucleotide was designated as the RP primer.

In order to obtain the full-length cDNA of pPAT5, 100 ng each of the RP primer and the oligo 2 primer, and 10 ng of pPAT5 digested with XhoI were subjected to PCR reaction in a final volume of 50  $\mu\text{l}$ . The reaction was carried out for 20 cycles with one cycle comprising 95°C for 1 minute, 55°C for 1 minute, and 72°C for 1 minute. The DNA fragment of about 600 bp thus obtained was run on an agarose gel electrophoresis and purified by GENECLEAN. After the fragment was enzyme digested with SmaI, the DNA fragment of about 400 bp was purified in a similar manner. The DNA fragment was labelled with the above-mentioned DIG.

Using the labelled DNA fragment, the above cDNA library of petals of petunias was screened by the plaque hybridization technique. Washing after hybridization was carried out in 0.2 x SSC at 65°C for 1 hour. Determination of the nucleotide sequence of the plasmid recovered from the clone obtained revealed that pPAT48 contained the same

sequence as pPAT5 does. This is shown in the sequence listing SEQ ID No. 3. This sequence had a homology of 20% and 16% with pGAT4 and pGAT106, respectively, at the level of amino acid sequence.

#### Example 9. Extraction of crude enzyme solution from perillas

Young red leaves were harvested from the plant bodies of *Perilla ocimoides* var. Akachirimren, and a crude enzyme solution was extracted according to the method as set forth in Example 1(2). This was reacted with 50  $\mu$ l of a mixture containing, at a final concentration, 50 mM potassium phosphate, pH 8.5, 0.48 mM delphinidin 3,5-diglucoside, 0.43 mM caffeoyl-CoA and 20  $\mu$ l of the enzyme solution at 30°C for 10 minutes. 50  $\mu$ l of acetonitrile containing 13.8% acetic acid was added to the reaction mixture to stop the reaction. After centrifuging at 15000 rpm for 5 minutes, a 10  $\mu$ l aliquot of the supernatant was analyzed by HPLC under the following conditions.

The column used was the YMC-Pack ODS-A (6.0 x 15 cm), and samples were separated under the condition of 0.1% trifluoroacetic acid, 21.6% acetonitrile, and a flow rate of 1 ml/min. Detection was carried out at 520 nm. Under this condition unreacted delphinidin 3,5-diglucoside was eluted at 3 minutes and the one in which caffeic acid was transferred to position 3 of delphinidin 3,5-diglucoside was eluted at 4.7 minutes, the absorption maximum of said compound being 531 nm.

Modification by caffeic acid was also seen when delphinidin 3-glucoside was used as the substrate. Furthermore, when coumaroyl-CoA was used as a donor of acyl group, transfer of a coumaroyl group was observed. It was revealed that although natural perillas do not contain delphinidin glucoside as anthocyanin, the acyltransferase of perillas can use delphinidin 3-glucoside and delphinidin 3,5-diglucoside as the acyl group recipient and coumaroyl-CoA as the acyl group donor.

#### Example 10. Purification of the acyltransferase derived from perillas

Purification of the acyltransferase derived from perillas was carried out in accordance with the method described in Example 2(1). Three kilograms of leaves of perillas was frozen in liquid nitrogen and pulverized frozen in a homogenizer. The pulverized material was homogenized again in 10 liters of the extraction buffer (100 mM sodium phosphate, pH 6.8, 10 mM sodium ascorbate, 5 mM dithiothreitol, 10  $\mu$ M p-APMSF, 5% (w/v) polyclar SB-100) in a homogenizer. This was filtrated with gauze stacked in four layers, and then centrifuged (8,000 rpm, 4°C, 30 minutes). Ammonium sulfate was added to the supernatant to a 40% saturation. After dissolution, centrifugation is repeated under the same condition. Ammonium sulfate was added to the supernatant to a 70% saturation. After dissolution, centrifugation is repeated under the same condition. The precipitate was dissolved in a minimum amount of the desalting buffer (bis Tris-HCl, pH 6.3, 1 mM dithiothreitol, 10  $\mu$ M p-APMSF, 10% glycerol), and then desalted by Sephadex G-25 medium (Pharmacia, 9.5 x 45 cm) which had been equilibrated with the same buffer.

The desalted sample was subjected to ion exchange chromatography using Q-Sepharose Fast Flow 26/10. A linear gradient of sodium chloride from 0 to 0.5 M in the desalting buffer was run at a flow rate of 8 ml/min over 1 hour. The active fractions were eluted at NaCl concentrations of about 0.15 to 0.3 M. The active fractions were adsorbed to four HiTrap Blue (5 ml) columns connected in a series which had been equilibrated with the desalting solution. After adequately washing the columns with the same buffer, elution was carried out by a linear gradient of sodium chloride from 0 to 1 M in the desalting buffer (2 hours, flow rate 5 ml/min). The active fractions were eluted at NaCl concentrations of 0.8 to 0.9 M. These fractions were subjected to chromatography using a hydroxyapatite column (ceramic type II 40 mm; Bio-Rad). The column on which a sample had been applied was adequately washed with buffer A (50 mM sodium phosphate, pH 6.8, 1 mM dithiothreitol, 10  $\mu$ M p-APMSF, 10% glycerol). Then enzyme was eluted with a liner gradient from buffer A to buffer B (400 mM sodium phosphate, pH 6.8, 1 mM dithiothreitol, 10  $\mu$ M p-APMSF, 10% glycerol), which eluted at about 0.2 M sodium phosphate. This active fractions were used for biochemical characterization of the enzyme.

In a similar manner to when the crude enzyme sample was used, any of cyanidin 3-glucoside, cyanidin 3,5-diglucoside, delphinidin 3-glucoside, and delphinidin 3,5-diglucoside could be used as the acyl group recipient. As the acyl group donor coumaroyl-CoA and caffeoyl-CoA could be used. The molecular weight was found to be about 50,000 by SDS-polyacrylamide gel electrophoresis. The isoelectric point was determined to be 5.3 using a Mono-P column (Pharmacia).

#### Example 11. cDNA cloning of the acyltransferase derived from perillas

By comparing the structures of pGAT4, pGAT106, and pGAT48 which were cloned in Example 3, Example 6, and Example 8, respectively, it was found that the amino acid sequence: Asp-Phe-Gly-Trp-Gly-Lys (SEQ ID No. 21) have been conserved. Accordingly, it is expected that this structure is also conserved in acyltransferases. Based on this sequence, nucleotide sequence: 5'-GA(TC)TT(TC)GGITGGGGIAA-3' (SEQ ID No. 22) was synthesized and this oligonucleotide was used as an ATC primer.

From young leaves of perillas, RNA was extracted in the method as described in Example 3 and a cDNA library was also constructed using the ZAP-cDNA synthesis kit (Stratagene). Using 50 ng of double stranded cDNA which was formed here as the template and 100 ng each of the ATC primer and the oligo 2 primer, PCR reaction was carried out in a final volume of 50  $\mu$ l using the PCR kit (Takara Shuzo). The reaction was carried out for 25 cycles with one cycle comprising 95°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute. A DNA fragment of about 400 bp thus obtained was recovered and cloned into a vector using the TA cloning kit (Invitrogen). Determination of the nucleotide sequence of the clone obtained revealed that the clone designated as pSAT104 had a high homology with pGAT4.

Using 10 ng of pSAT104 as the template and 100 ng each of the ATC primer and the oligo 2 primer, PCR reaction was carried out using the PCR kit (Takara Shuzo K.K.) in a final volume of 50  $\mu$ l. The reaction was carried out for 15 cycles with one cycle comprising 95°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute. Using 1  $\mu$ l of this product and 100 ng each of the ATC primer and the oligo 2 primer, PCR reaction was carried out using the PCR kit in a final volume of 50  $\mu$ l.

However, 4  $\mu$ l of the DIG-labelled nucleotide solution (manufactured by Boehringer) was used here as the deoxy-nucleotide solution. After the reaction was complete, 5  $\mu$ l of 3 M sodium acetate and 100  $\mu$ l of ethanol were added to carry out ethanol precipitation. The product obtained was used for the subsequent study.

Using the labelled DNA fragment derived from pSAT104, the cDNA library of leaves of perillas was screened by the plaque hybridization technique. Washing was carried out in 1 x SSC at 65°C for 1 hour. Determination of the nucleotide sequence of the hybridized clone revealed that the clones of pSAT206, pSAT207, pSAT208, pSAT209, pSAT210, etc. contains the nucleotide sequence of pSAT104. When the nucleotide sequence of the 5' ends of these clones were compared with that of pGAT4, every clone had an amino terminal shorter than pGAT4 and none had the initiation codon. The nucleotide sequences of 5' ends of pSAT206 and pSAT208, and pSAT209 and pSAT210 were identical. pSAT207 was shorter than pSAT206 by 6 residues, and pSAT209 was shorter than pSAT206 by 5 residues.

On the vector pBluescript SK-, these cDNA's are taking such forms that enable them to fuse to the LacZ gene of the vector. Out of the above-mentioned clones, pSAT206, pSAT208, and pSAT207 are taking such shapes that enable them to express as a fusion protein with  $\beta$ -galactosidase, whereas pSAT209 and pSAT210 have shifted frames so that they cannot form a fusion protein. pSAT206, pSAT207, pSAT209, and pSAT210 were expressed in *E. coli*, and then tested for the enzymatic activity of acyl group transfer to position 3 of glucose using delphinidin 3,5-diglucoside and coenzyme-CoA. The method for inducing expression etc. was carried out in accordance with the method as set forth in Example 4.

The *E. coli*'s containing pSAT209 and pSAT210 did not exhibit any enzymatic activity of transferring acyl groups, but the *E. coli* containing pSAT206 exhibited the enzymatic activity of acylating 48% of delphinidin 3,5-diglucoside and the *E. coli* containing pSAT207 exhibited a similar enzymatic activity of acylating 24% of said compound. These results demonstrated that pSAT206, pSAT207, and the like reveal cloning of the gene having the enzymatic activity of transferring acyl groups to glucose at position 3 of anthocyanin of perillas.

Among these clones, the nucleotide sequence derived from cDNA of pSAT208 was determined. This is shown in the sequence listing SEQ ID No. 4. The amino acid sequence deduced from the nucleotide sequence exhibited a homology of 37%, 29%, and 15% with pGAT4, pGAT106, and pPAT48, respectively. As described hereinbefore, this sequence, though not a full-length cDNA, can express active enzymes by providing a suitable initiation codon as a fusion gene with LacZ.

By comparing the amino acid sequences of acyltransferases which were elucidated by the present invention, the conserved sequence was clarified. Based on the amino acid sequence of this region, it is possible to clone acyltransferases which modify sugars at other positions of anthocyanins.

#### Example 12. cDNA cloning of the acyltransferase derived from cinerarias

From petals of *Senecio cruentus* var. Jupiter Blue (Sakata Seed Corp.), RNA was extracted by the method as set forth in Example 3 above and Poly A+ RNA was further purified. A cDNA library was constructed using the ZAP-cDNA synthesis kit (Stratagene).

Using 50 ng of double stranded cDNA which was formed here as the template and 100 ng each of the ATC primer and the oligo 2 primer, PCR reaction was carried out in a final volume of 50  $\mu$ l using the PCR kit (Takara Shuzo). The reaction was carried out for 25 cycles with one cycle comprising 95°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute. DNA fragments of about 400 bp thus obtained were collected and cloned into a vector using the TA cloning kit (Invitrogen). Determination of the nucleotide sequence of clones obtained revealed that a clone designated as pJAT4 had a high homology with pGAT4.

Then the cDNA library of petals of cinerarias was screened with pJAT4. Several clones were obtained. When the amino acid sequences deduced from the nucleotide sequences of the 5' end of these cDNA's were compared with the sequence of the protein encoding pGAT4, none of cDNA of the clones of cinerarias were full-length. Among them the entire nucleotide sequence of cDNA of the clone termed pCAT8 was determined. This is shown in the sequence listing SEQ ID No. 5. The amino acid sequence deduced from the nucleotide sequence obtained exhibited a homology of

28%, 35%, 16%, and 37% with pGAT4, pGAT106, pPAT48, and pSAT208, respectively.

Example 13. Construction of a binary vector containing the gene of the acyltransferase derived from gentian

- After the acyltransferase gene of gentians, pGAT4, was completely digested with KpnI, a DNA fragment of about 1.6 kb which is obtained by partial digestion thereof with XbaI was collected. This DNA fragment was subcloned using the restriction enzyme recognition sites, KpnI and XbaI, of pUC19 to obtain a plasmid pUCGAT4. After pUCGAT4 was completely digested with BglII, it was partially digested with SacI to collect a DNA fragment of about 0.95 kb. A plasmid obtained by ligating this DNA fragment, about 0.75 kb DNA fragment obtained by digestion of pUCGAT4 with XbaI and BglII, and a DNA fragment obtained by digestion of plasmid p2113G (described, for example, in Aida et al., *Acta Horticulture*, 392: 219-225, 1995) with XbaI and SacI was designated as pBEGA4. This plasmid is a binary vector, and the gentian acyl transferase cDNA is under the control of the cauliflower mosaic virus 35S promoter having enhancers and the nopaline synthase terminator within plant cells. It also has a translation enhancer called  $\Omega$  sequence at the 5' end of cDNA of the gentian acyltransferase. It is noted that the promoter and terminator as used herein are not limited to those just described, but they may be a constitutive promoter or a promoter which specifically works in petals such as the promoter of the gene of chalcone synthase.

Example 14. Introduction of the acyltransferase gene derived from gentian into plants

- pBEGA4 was introduced into *Agrobacterium tumefaciens* strain Ag10 (Lazo et al., *Bio/Technology*, 9: 963-967, 1991) by the method described in *Plant Molecular Biology Manual* (Kluwer Academic Publishers). On the other hand, by culturing a shoot apex of a rose var. Lavende in a solid medium in which BA (6-benzyl aminopurine) 2.25 mg/l, GA3 (gibberellic acid) 3.46 mg/l, sucrose 30 g/l, and Gelatin Gum 2 g/l were added to the MS medium to obtain Embryogenic Callus (EC). An overnight culture of the above AG10 strain in the LB medium was suspended to the MS liquid medium containing 20  $\mu$ g/ml of acetosyringone to adjust to a concentration of  $5 \times 10^8$  cells/ml. After immersing the EC in this bacterial culture liquid, excess liquid was wiped clean by sterilized filter paper. By transplanting and culturing in the MS medium in which BA 2.25 mg/l, GA3 0.35 mg/l, sucrose 30 g/l, and Gelatin Gum 2 g/l were added to the MS medium, a transformant can be obtained. From the kanamycin resistant callus obtained, RNA was obtained using trizol (Life/tec Oriental). Using this RNA as the template, and nucleotide GAT-1: 5'-TGGCAACTGTCTTCGCTCATG-3' (SEQ ID No. 23) and nucleotide GAT-2: 5'-CCATGTCAGGTGTGAGGTTCAAC-3' (SEQ ID No. 24) synthesized based on the nucleotide sequence of pGAT4 as the primer, RT-PCR reaction was carried out using the Access RT-PCR System (Promega). Using the same RNA as the template, and oligonucleotide Kan-1: 5'-ATCGTTTCGCATGATTGAAC-3' (SEQ ID No. 25) and oligonucleotide Kan-2: 5'-TCAGAAGAAGCTCGTCAAGAA-3' (SEQ ID No. 26) synthesized based on the nucleotide sequence of nptII on the binary vector as the primer, RT-PCR reaction was similarly carried out. The reaction was carried out for 40 cycles with one cycle comprising 94°C for 30 seconds, 60°C for 1 minute, and 68°C for 2 minutes. From the callus of the transformant a band derived from pGAT4 and a band derived from nptII were observed, but from the callus of the non-transformant no bands corresponding to the above were observed. The result indicates that the gene of acyltransferase of gentians could be introduced into the rose.

- Construction of the binary vector mentioned above and its transformation into a plant are not limited to the gene of acyltransferase of gentians contained in pGAT4, but other acyltransferases can be introduced into plants and genes thereof can be expressed in plants. As a species of a plant, a rose was described hereinabove. But since methods of transformation have been reported for many other plants (for example, carnations, chrysanthemums, tobaccos, petunias, gerberas, petunias, etc.), acyltransferase could be introduced into many plant species by employing published methods.

Example 15. Synthesis of the full-length cDNA of the acyltransferase derived from perillas

- The cDNA of acyltransferase gene of perillas, pSAT208, encodes active enzymes as described above, but it was not full-length. Accordingly, a full-length cDNA containing the initiation codon was synthesized based on the nucleotide sequence of acyltransferase gene of gentians, pGAT4. Thus, the DNA shown below was synthesized. The amino acid sequence encoded by the DNA is also shown. The first underline means a BamHI recognition sequence, and the next underline means a sequence contained in pSAT208. Behind the BamHI recognition sequence is inserted a sequence AACA which often occurs immediately before the translation initiation codon in plants.

5' GGATCCAACA ATG GAG CAA ATC CAA ATG GTG GCC GTG ATC GAA ACG TGT AGA 3'

Met Glu Gln Ile Gln Met Val Ala Val Ile Glu Thr Cys Arg

(SEQ ID No. 27)

PCR reaction was carried out in a final volume of 50  $\mu$ l containing 100 ng each of this primer and -20 primer: 5'-GTAAACGACGGCCAT-3' (SEQ ID No. 28), and 10 ng of pSAT208. The reaction was carried out for 15 cycles with one cycle comprising 95°C for 1 minute, 55°C for 1 minute, and 72°C for 2 minutes. After the reaction, DNA fragments were recovered from the reaction mixture using GENECLAN (BioL01) by the method recommended by the manufacturer. After digesting the recovered DNA with BamHI and EcoRI, followed by agarose gel electrophoresis, a DNA fragment of about 200 bp was recovered. This DNA fragment was ligated to a DNA fragment of about 3.3kb obtained by digesting pSAT208 with EcoRI and the plasmid obtained was termed pSATF208. The nucleotide sequence of this plasmid was determined from 5' end of the cDNA to confirm the nucleotide sequence.

#### Example 16. Expression of the gene of the acyltransferase derived from perillas in yeast

In accordance with the method described in Example 5, pSATF208 was expressed in yeast and tested for enzymatic activity. Thus, a plasmid obtained by ligating a DNA fragment of about 8 kb obtained by digestion of pYE22m with BamHI and SalI with a DNA fragment of about 8 kb obtained by digestion of pSATF208 with BamHI and SalI was termed pYSAT208.

Yeast G1315 was transformed with pYSAT208 and the activity of acyltransferase of the resulting transformant was determined. As a result, in the yeast into which pYSAT208 was introduced, formation of 10 nmole of delphinidin 3-cafeoylglycoside 5-glucoside from 24 nmol of delphinidin 3,5-diglucoside and 21.5 nmol of caffeoyl-CoA was observed. Thus, it was confirmed that the synthesized full-length cDNA contained in pSATF208 encodes the activity of acyltransferase.

#### Example 17. Construction of a binary vector containing the acyltransferase gene derived from perillas

Plasmid pE12 $\alpha$ GUS is one in which the expression unit of GUS gene on plasmid p2113G (Aida et al., Acta Horticulture, 392: 219-225, 1995) has been inserted into the HindIII and EcoRI recognition sites of pUC19. After pE12 $\alpha$ GUS was digested with SacI and blunt-ended using the DNA blunting kit (Takara Shuzo), it was ligated to an XhoI linker (Toyobo). The plasmid obtained which has an XhoI linker inserted thereinto was termed pE12 $\alpha$ GUSx. A plasmid obtained by ligating about 2.8 kb DNA fragment obtained by digestion of this plasmid with HindIII and EcoRI to pBin19 digested with HindIII and EcoRI was designated pBEGUSx. A plasmid obtained by ligating a DNA fragment of about 11 kb obtained by digestion of BEGUSx with BamHI and XhoI to a DNA fragment obtained by digestion of pSATF208 with BamHI and XhoI was termed pBESA208. On this plasmid the acyltransferase of perillas is under the control of the cauliflower mosaic virus 35S promoter having enhancers and the nopaline synthase terminator.

#### Example 18. Introduction of the acyltransferase gene of perillas into plants

pBESA208 was introduced into Agrobacterium tumefaciens strain Ag10 (Lazo et al., Bio/Technology, 9: 963-967, 1991) by the method described in Plant Molecular Biology Manual (Kluwer Academic Publishers). Ag10 strain transformant was used to transform petunia Falcon red (Sakata Seed Co.), Baccarat red (Sakata Seed Co.), and Titan red (Sakata Seed Co.) by the method described in Plant Molecular Biology Manual (Kluwer Academic Publishers). These petals of petunia contain cyanidin-3-glucoside as a major anthocyanin.

Transformation was also carried out to a rose var. Lavande by the above-mentioned method.

#### Example 19. Synthesis of the full-length cDNA of the acyltransferase derived from cinerarias

The cDNA of cineraria acyltransferase gene, pCAT208, as hereinabove described, was not full-length. Accordingly, a full-length cDNA containing the initiation codon was synthesized based on the nucleotide sequence of the gentian acyltransferase gene, pGAT4. Thus, the DNA shown below was synthesized. The amino acid sequence encoded by the DNA is also shown. The first underline means a BamHI recognition sequence for cloning, and the next underline means a sequence contained in pCAT208.



5' GGGATCCAACA ATG GAG CAA ATC CAA ATG GTG AAC ATT CTC GAA C 3'

Met Glu Gln Ile Gln Met Val Asn Ile Leu Glu

(SEQ ID No. 29)

PCR reaction was carried out in a final volume of 50  $\mu$ l containing 100 ng each of this primer and -20 primer, and 10 ng of pCAT8. The reaction was carried out for 15 cycles with one cycle comprising 95°C for 1 minute, 55°C for 1 minute, and 72°C for 2 minutes. After the reaction was over, DNA fragments were recovered from the reaction mixture using GENECLAN (Bio101) by the method recommended by the manufacturer. After digesting the recovered DNA with BamHI and MvaI, followed by agarose gel electrophoresis, a DNA fragment of about 200 bp was recovered. This DNA fragment was ligated to a DNA fragment of about 1.3kb obtained by digesting pCAT8 with MvaI and XhoI and plasmid pBluescript SK- digested with BamHI and XhoI, and the plasmid obtained was termed pCATF208. The nucleotide sequence of this plasmid was determined from the 5' end of the cDNA to confirm the nucleotide sequence.

#### Example 20. Cloning of cDNA encoding the acyltransferase derived from lavenders

A cDNA library derived from petals of lavender of the perilla family, *Lavandula angustifolia*, was constructed by the method as set forth in Example 3, and screened by the plaque hybridization technique detailed in Example 3. About 300,000 clones were screened. Thus, the probe used was obtained by carrying out PCR reaction in a final volume of 50  $\mu$ l using 100 ng each of the synthetic nucleotide RI primer: 5'-CTCGGAGGAATTCGGCAGCAG-3' (SEQ ID No. 30) and the oligo 2, 10 ng of pSAT208, and a DIG-labelled nucleotide as the nucleotide. The reaction was carried out for 25 cycles with one cycle comprising 95°C for 1 minute, 50°C for 1 minute, and 72°C for 2 minutes.

The labelled cDNA fragment was added to the hybridization solution and hybridization was carried out at 37°C for further 16 hours. The filter was washed with the washing solution (5 x SSC, 1% SDS) and then an enzymeimmunoassay (Boehringer Mannheim) using the DIG-specific antibody labelled with alkaline phosphatase was carried out to detect positive clones by color development using 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium salt. The detection method used was as set forth in the manufacturer's instructions.

As a result, one positive clone was obtained. This cDNA was rescued in the form which employs plasmid pBlue-script SK- as a vector from the form which employs  $\lambda$  phage as the vector using the method recommended by the manufacturer. A plasmid was extracted from the clone obtained and was termed pLAT1. As described above, the nucleotide sequence in the vicinity of the 5' end of cDNA of pLAT1 was determined using the ABI373 DNA Sequencer (Perkin Elmer) by the pigment deoxy sequence method with the fluorogenic pigment recommended by the manufacturer. The amino acid sequence deduced from the nucleotide sequence thus obtained has a high homology with the amino acid sequence of acyltransferase of perillas and gentians, suggesting that pLAT1 encodes the acyltransferase of lavenders. But, the amino acid sequence encoded by pLAT1 is shorter than that of acyltransferase of perillas or gentians and it is, conceivably, not long enough to encode the entire length of acyltransferase. Accordingly, using the cDNA fragment of pLAT1 labelled with DIG as the probe, the cDNA library of lavenders was screened under the same condition mentioned above. The probe was labelled by the PCR reaction in a final volume of 50  $\mu$ l containing about 1 ng of pLAT1 plasmid as the template, 500 ng each of the RI primer described below and the oligo 2, and 8  $\mu$ l of dNTP-labelled mixture (Boehringer). The PCR reaction was carried out for 25 cycles with one cycle comprising 95°C for 1 minute, 42°C for 2 minutes, and 72°C for 3 minutes, and the reaction was kept at 72°C for 7 more minutes in order to perfect the elongation reaction. The plaque hybridization was carried out by the method described above except that the concentration of formamide in the hybridization buffer was 50% and that the filter was washed with 2 x SSC and 1% SDS. The nucleotide sequence in the vicinity of the 5' end of cDNA of the positive clone obtained was determined as described above, and a clone, pLAT21, which is 11 bp longer than pLAT1 was obtained. This is shown in the sequence listing SEQ ID No. 6. However, pLAT21 did not contain the methionine initiation codon either and was not long enough to encode the entire length.

#### Example 21. Synthesis of the full-length cDNA of the acyltransferase derived from lavenders

Since the cDNA, pLAT21, which is considered to encode the acyltransferase of lavenders does not contain the methionine initiation codon, the methionine initiation codon must be added to the 5' end of the cDNA in order to permit its expression in yeast. Accordingly, using a primer as described below, PCR reaction was carried out to synthesize a fragment in which the methionine initiation codon has been added to the 5' end of pLAT21. The primer LAT-ATG is designed so that it contains, in addition to 20 nucleotide sequences at the 5' end of pLAT21, the methionine initiation

codon, the conserved sequence AACA for gene expression in plant which is believed to be present adjacent to the upstream thereof and the restriction enzyme BamHI recognition site required for ligation to a yeast expression vector in the direction of 5' upstream to 3'. The LAT-ATG primer (SEQ ID No. 31):

5' -AGTCGGATCCAACA ATG ACC ACC CTC CTC GAA TCC 3'

Thr Thr Leu Leu Glu Ser

PCR reaction was carried out in a final volume of 50 µl containing about 100 ng of the pLAT21 plasmid as the template, and 500 ng each of the LAT-ATG primer and the oligo 2 primer. PCR reaction was carried out for 10 cycles with one cycle comprising 95°C for 1 minute, 42°C for 2 minutes, and 72°C for 3 minutes, and the reaction was kept at 72°C for 7 more minutes in order to perfect the elongation reaction. The DNA fragment thus obtained was cleaved with BamHI and EcoRI and a DNA fragment of about 550 bp was recovered. This DNA fragment was subcloned into the BamHI and EcoRI sites of the plasmid vector plasmid pBluescript SK- and termed pLATPCR11. The nucleotide sequence of pLATPCR11 was determined as mentioned before, and it was confirmed that this PCR-amplified DNA fragment had the same sequence as that of from the 5' end to EcoRI site of pLAT21 cDNA, and contained the methionine initiation codon in the LAT-ATG primer and the conserved sequence for gene expression in plants, and the restriction enzyme BamHI recognition site required for ligation to a yeast expression vector.

Furthermore, the entire nucleotide sequence of pLAT21 was determined in a similar method to the one used to determine the nucleotide sequence of cDNA of pGAT4. The amino acid sequence expected to be encoded by this cDNA had a homology of 69%, 38%, 37%, 37%, and 19% with pSAT208, pGAT4, pGAT8, pGAT106, and pPAT48, respectively.

#### Example 22. Expression of the acyltransferase gene derived from lavenders in yeast

A plasmid obtained by ligating a DNA fragment of about 550 bp cleaved out from pLATPCR11 with BamHI and EcoRI, a DNA fragment of about 1 kb obtained by cleavage of pLAT21 with EcoRI and XhoI, and a DNA fragment of about 8 kb obtained by cleavage of pYE22m with BamHI and SalI was designated as pYELAT21. As hereinabove explained, yeast G1315 was transformed with this plasmid and the activity of acyltransferase was determined.

As a result, in the yeast into which pYELAT21 was introduced, formation of 19.9 nmol of delphinidin 3-caffeoylglucoside 5-glucoside from 24 nmol of delphinidin 3,5-diglucoside and 21.5 nmol of caffeoyl-CoA was observed. Thus, it was confirmed that the synthesized full-length cDNA contained in pYELAT21 encodes the activity of acyltransferase.

#### Example 23. Construction of a binary vector containing the acyltransferase gene derived from lavenders

A plasmid obtained by ligating a DNA fragment of about 550 bp cleaved out from pLATPCR11 with BamHI and EcoRI, a DNA fragment of about 1 kb obtained by cleavage of pLAT21 with EcoRI and XhoI, and a DNA fragment of about 11 kb obtained by digestion of pBEGUSx with BamHI and XhoI was designated as pBELA11. As hereinabove explained, this was transformed into Agrobacterium tumefaciens strain Ag1 0 and was supplied for transformation of petunias and roses.

#### Example 24. Construction of antibody against acyltransferase

As a means to obtain the gene of an enzyme whose amino acid sequence is similar to that of the desired enzyme, there is mentioned a method in which the cDNA library of the expression form is screened by antibody to an enzyme. In this case, an antibody against the acyltransferase encoded by pGAT4 of gentians was produced. Similarly, it is possible to produce antibodies to other acyltransferases.

First, using the Bulk and RediPack GST Purification Modules (pharmacia Biotech), the GAT4 protein was expressed in large quantities using *E. coli*, from which the antibody was purified.

#### (1) Construction of expression plasmid

pGEX-4T-1 was used to express the acyltransferase gene in *E. coli*. Using this pGEX-4T-1 a fusion protein with glutathione S-transferase can be prepared, which is purified efficiently using an affinity column of glutathione S-transferase.

After pGEX-4T-1 was digested with SmaI and XhoI, it was blunt-ended using the DNA blunting kit (Takara Shuzo). The DNA fragment of about 4.9 kb thus obtained was dephosphorylated using alkaline phosphatase BAP C75 (Takara

Shuzo). A DNA fragment of about 1.6 kb obtained by digestion of pGAT4 with SmaI and KpnI present in said vector was blunt-ended as described before, and was recombined with the above-mentioned blunt-ended site after digestion of pGEX-4T-1 with SmaI and XhoI to construct pGEXGAT4. By digesting with EcoRI and BglII, it was confirmed that the cDNA and glutathione S-transferase on pGAT4 were in the same direction.

## (2) Expression of acyltransferase in *E. coli*

*E. coli* strain JM109 was transformed with pGEXGAT4. Transformation of *E. coli* was carried out by the method of Hanahan (J. Mol. Biol. 166: 557-, 1983). The transformed *E. coli* was inoculated to 50 ml of 2 x YT medium (tryptone 16 g, yeast extract 10 g, and sodium chloride 5 g were dissolved into one liter of distilled water, and then pH was adjusted to 7.0 with sodium hydroxide) containing ampicillin (100 µg/l) and 2% glucose, and then cultured overnight at 37°C. Forty ml of the culture was inoculated into 40 ml of 2 x YT medium containing ampicillin (100 µg/l) and 2% glucose, followed by incubation at 37°C for 3 hours, to which 440 µl of IPTG (final concentration 10 mM) was added and cultured for 5 more hours. After harvesting, the cells were suspended to 100 ml of 1 x PBS (sodium chloride 8.2 g, potassium chloride 2.0 g, disodium hydrogen phosphate 1.43 g, and potassium dihydrogen phosphate 2.45 g were dissolved in one liter of distilled water) containing 10 µM of APMSF. After the suspension was disrupted by sonication, 5 ml of 20% Triton X-100 was added (final concentration 1%). After shaking in ice for 30 minutes, it was centrifuged at 12,000 rpm for 10 minutes. The precipitate obtained was suspended to 12 ml of 6 M urea, to which was added an equal amount of 2 x SDS sample buffer and treated at 90°C for 5 minutes to prepare a sample.

This sample (0.8 ml) was separated on disk gel electrophoresis (separating gel 7.5% acrylamide, stacking gel 5% acrylamide; ATTO BIO PHORESIS I II) and collected in aliquots of 0.8 ml. Each fraction was analyzed on SDS-polyacrylamide gel electrophoresis (separating gel 10% acrylamide, stacking gel 4.5% acrylamide). The result indicated that there was a fraction in which a protein having a molecular weight of about 75,000 corresponding to the size of a fusion protein of the acyltransferase and glutathione S-transferase encoded by pGAT4 was present as a single protein.

This fraction (3.2 ml) was concentrated by Centricon 10 (Amicon) to obtain about 0.3 µg of the fusion protein. Using this sample, antibody was produced using BALB/C mice by the conventional method. Using this antibody, a homologs of acyltransferase can be obtained. Industrial Applicability

As hereinabove described, in accordance with the present invention, aromatic acyltransferase derived from gentians was purified, the cDNA of said enzyme was cloned, and the nucleotide sequence of said cDNA was determined.

Furthermore, by expressing the activity in *E. coli* and yeast, the separated cDNA was confirmed to be identical with the one encoding aromatic acyltransferase.

Thus, by connecting the cDNA according to the present invention to a suitable plant expression vector and then introducing it into a plant, it became possible to utilize acylation reaction in order to control the color of flowers.

Furthermore, by utilizing the present enzymatic activity, it is possible to modify the structures of anthocyanins in plants or in vitro in order to provide more stable anthocyanins.

## SEQUENCE LISTING

SEQ ID NO: 1

SEQUENCE LENGTH: 1703

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL SEQUENCE: No

ANTI-SENSE: No

ORIGINAL SOURCE

ORGANISM: *Gentiana triflora* var. *japonica*

TISSUE TYPE: petal

IMMEDIATE SOURCE

LIBRARY: cDNA library

CLONE: pGAT4

SEQUENCE DESCRIPTION

TCATT ATG GAG CAA ATC CAA ATG GTG AAG GTT CTT GAA AAA TGC CAA	47
Met Glu Gln Ile Gln Met Val Lys Val Leu Glu Lys Cys Gln	
-1 1 5 10	
GTT ACA CCA CCA TCT GAC ACA ACA GAT GTC GAG TTA TCG CTA CCG GTA	95
Val Thr Pro Pro Ser Asp Thr Thr Asp Val Glu Leu Ser Leu Pro Val	
15 20 25	
ACA TTC TTC GAT ATC CCC TGG TTG CAC TTG AAT AAG ATG CAG TCC CTT	143
Thr Phe Phe Asp Ile Pro Trp Leu His Leu Asn Lys Met Gln Ser Leu	
30 35 40 45	
CTG TTT TAC GAC TTT CCG TAC CCA AGA ACA CAT TTC TTG GAC ACT GTT	191
Leu Phe Tyr Asp Phe Pro Tyr Pro Arg Thr His Phe Leu Asp Thr Val	
50 55 60	
ATC CCT AAT CTT AAG GCC TCT TTG TCT CTC ACT CTA AAA CAC TAC GTT	239
Ile Pro Asn Leu Lys Ala Ser Leu Ser Leu Thr Leu Lys His Tyr Val	
65 70 75	
CCG CTT AGC GGA AAT TTG TTG ATG CCG ATC AAA TCG GGC GAA ATG CCG	287
Pro Leu Ser Gly Asn Leu Leu Met Pro Ile Lys Ser Gly Glu Met Pro	
80 85 90	
AAG TTT CAG TAC TCC CGT GAT GAG GGC GAC TCG ATA ACT TTG ATC GTT	335
Lys Phe Gln Tyr Ser Arg Asp Glu Gly Asp Ser Ile Thr Leu Ile Val	
95 100 105	

EP 0 810 287 A1

	GCG GAG TCT GAC CAG GAT TTT GAC TAC CTT AAA GGT CAT CAA CTG GTA	383
5	Ala Glu Ser Asp Gln Asp Phe Asp Tyr Leu Lys Gly His Gln Leu Val	
	110 115 120 125	
	GAT TCC AAT GAT TTG CAT GGC CTT TTT TAT GTT ATG CCA CGG GTT ATA	431
	Asp Ser Asn Asp Leu His Gly Leu Phe Tyr Val Met Pro Arg Val Ile	
10	130 135 140	
	AGG ACC ATG CAA GAC TAT AAA GTG ATC CCG CTC GTA GCC GTG CAA GTA	479
	Arg Thr Met Gln Asp Tyr Lys Val Ile Pro Leu Val Ala Val Gln Val	
	145 150 155	
15	ACC GTT TTT CCT AAC CGT GGC ATA GCC GTG GCT CTG ACG GCA CAT CAT	527
	Thr Val Phe Pro Asn Arg Gly Ile Ala Val Ala Leu Thr Ala His His	
	160 165 170	
20	TCA ATT GCA GAT GCT AAA AGT TTT GTA ATG TTC ATC AAT GCT TGG GCC	575
	Ser Ile Ala Asp Ala Lys Ser Phe Val Met Phe Ile Asn Ala Trp Ala	
	175 180 185	
	TAT ATT AAC AAA TTT GGG AAA GAC GCG GAC TTG TTG TCC GCG AAT CTT	623
25	Tyr Ile Asn Lys Phe Gly Lys Asp Ala Asp Leu Leu Ser Ala Asn Leu	
	190 195 200 205	
	CTT CCA TCT TTC GAT AGA TCG ATA ATC AAA GAT CTG TAT GGC CTA GAG	671
	Leu Pro Ser Phe Asp Arg Ser Ile Ile Lys Asp Leu Tyr Gly Leu Glu	
30	210 215 220	
	GAA ACA TTT TGG AAC GAA ATG CAA GAT GTT CTT GAA ATG TTC TCT AGA	719
	Glu Thr Phe Trp Asn Glu Met Gln Asp Val Leu Glu Met Phe Ser Arg	
35	225 230 235	
	TTT GGA AGC AAA CCC CCT CGA TTC AAC AAG GTA CGA GCT ACA TAT GTC	767
	Phe Gly Ser Lys Pro Pro Arg Phe Asn Lys Val Arg Ala Thr Tyr Val	
	240 245 250	
40	CTC TCC CTT GCT GAA ATC CAG AAG CTA AAG AAC AAA GTA CTG AAT CTC	815
	Leu Ser Leu Ala Glu Ile Gln Lys Leu Lys Asn Lys Val Leu Asn Leu	
	255 260 265	
45	AGA GGA TCC GAA CCG ACA ATA CGT GTA ACG ACG TTC ACA ATG ACG TGT	863
	Arg Gly Ser Glu Pro Thr Ile Arg Val Thr Phe Thr Met Thr Cys	
	270 275 280 285	
	GGA TAC GTA TGG ACA TGC ATG GTC AAA TCA AAA GAT GAC GTC GTA TCA	911
50	Gly Tyr Val Trp Thr Cys Met Val Lys Ser Lys Asp Asp Val Val Ser	
	290 295 300	

EP0810287A1

GAG GAA TCA TCG AAC GAC GAA AAT GAG CTC GAG TAC TTC AGT TTT ACA 959  
 Glu Glu Ser Ser Asn Asp Glu Asn Glu Leu Glu Tyr Phe Ser Phe Thr  
 5 305 310 315  
 GCG GAT TGC CGA GGA CTT CTG ACG CCC CGG TGT CCG CCT AAC TAC TTT 1007  
 Ala Asp Cys Arg Gly Leu Leu Thr Pro Pro Cys Pro Pro Asn Tyr Phe  
 320 325 330  
 10 GGC AAC TGT CTT GCG TCA TGC GTT GCA AAA GCA ACA CAT AAA GAG TTA 1055  
 Gly Asn Cys Leu Ala Ser Cys Val Ala Lys Ala Thr His Lys Glu Leu  
 335 340 345  
 15 GTT GGG GAT AAA GGG CTT CTT GTT GCA GTT GCA GCT ATT GGA GAA GCC 1103  
 Val Gly Asp Lys Gly Leu Leu Val Ala Val Ala Ala Ile Gly Glu Ala  
 350 355 360 365  
 ATT GAA AAG AGG TTG CAC AAC GAA AAA GGC GTT CTT GCA GAT GCA AAA 1151  
 Ile Glu Lys Arg Leu His Asn Glu Lys Gly Val Leu Ala Asp Ala Lys  
 20 370 375 380  
 ACT TGG TTA TCG GAA TCT AAT GGA ATC CCT TCA AAA AGA TTT CTC GGG 1199  
 Thr Trp Leu Ser Glu Ser Asn Gly Ile Pro Ser Lys Arg Phe Leu Gly  
 25 385 390 395  
 ATT ACC GGA TCG CCT AAG TTC GAT TCG TAT CGT GTA GAT TTT GGA TGG 1247  
 Ile Thr Gly Ser Pro Lys Phe Asp Ser Tyr Gly Val Asp Phe Gly Trp  
 30 400 405 410  
 GGA AAG CCT GCA AAA TTT GAC ATT ACC TCT GTT GAT TAT GCA GAA TTG 1295  
 Gly Lys Pro Ala Lys Phe Asp Ile Thr Ser Val Asp Tyr Ala Glu Leu  
 415 420 425  
 35 ATT TAT GTG ATT CAG TCC AGG GAT TTT GAA AAA GGT GTG GAG ATT GGA 1343  
 Ile Tyr Val Ile Gln Ser Arg Asp Phe Glu Lys Gly Val Glu Ile Gly  
 430 435 440 445  
 40 GTA TCA TTG CCT AAG ATT CAT ATG GAT GCA TTT GCA AAA ATC TTT GAA 1391  
 Val Ser Leu Pro Lys Ile His Met Asp Ala Phe Ala Lys Ile Phe Glu  
 450 455 460  
 GAA GGC TTT TGC TCT TTG TCA TAGTCTCTTT AATAGAACCA TATTGCTGCG 1442  
 Glu Gly Phe Cys Ser Leu Ser  
 465 468  
 AATAAAGTAC CAAAGTCCTTT AGTAACACTA CACCAAAACCC TACTTTCGAG GCGGGAACAC 1502  
 CACAACGAGG TTCAATCACT AGAAGGTTGT ACTTCATAAA TTCCAGAGGT CGAATATACA 1562  
 50 CCGTGTCTCT CTGAAAAGTT GAACCTCACA CCTGACATGG TGTTACGATA GGTATTGTAT 1622  
 AATGCCATTA TATACTTCCA TAAAGTATCC TATGCAATAG AGAACATGTT ATGTGTTAAA 1682

55

AAAAAAAAAA AAAAAAAAAA A 1703  
 SEQ ID NO: 2  
 SEQUENCE LENGTH: 1622  
 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL SEQUENCE: No  
 ANTI-SENSE: No  
 ORIGINAL SOURCE  
 ORGANISM: *Gentiana triflora* var. *japonica*  
 TISSUE TYPE: petal  
 IMMEDIATE SOURCE  
 LIBRARY: cDNA library  
 CLONE: pGAT106  
 SEQUENCE DESCRIPTION  
 GAACCATTTGA ATCCAATTAA TCTGATTAT TAAG ATG GCA GGA AAT TCC GAG 52  
 Met Ala Gly Asn Ser Glu  
 1 5  
 GAT ATC AAA GTT CTT GAG AAA TGC CGT GTT GCG CCA CCA CCG GAC GCC 100  
 Asp Ile Lys Val Leu Glu Lys Cys Arg Val Ala Pro Pro Pro Asp Ala  
 10 15 20  
 GTC GCC GAG TTT ACA GTC CCA CTG TCG TTT TTC GAC ATG CGA TGG TTG 148  
 Val Ala Glu Phe Thr Val Pro Leu Ser Phe Phe Asp Met Arg Trp Leu  
 25 30 35  
 ATC TCT GAT GCA GAA CAC CAT CTG CAT TTC TAC AGA TTC CGC CAT CCT 196  
 Ile Ser Asp Ala Glu His His Leu His Phe Tyr Arg Phe Arg His Pro  
 40 45 50  
 TGT CCC AAC TCT AAA TTT ATC ATT TCA TCC ATT AAA TCG TCC CTT TCC 244  
 Cys Pro Asn Ser Lys Phe Ile Ile Ser Ser Ile Lys Ser Ser Leu Ser  
 55 60 65 70  
 CTT GTT CTC AAA CAC TTT CTT CCG TTA GCC GGG AAT TTG ATT TGG CCG 292  
 Leu Val Leu Lys His Phe Leu Pro Leu Ala Gly Asn Leu Ile Trp Pro  
 75 80 85  
 GTA GAT TCC TCC GAT AGA ATG CCG GAG TTG CGT TAC AAG AAA GGC GAC 340  
 Val Asp Ser Ser Asp Arg Met Pro Glu Leu Arg Tyr Lys Lys Gly Asp  
 90 95 100

EP 0 810 287 A1

	TCC GTT TCT TTA ACA ATT GCA GAA TCG AGC ATG GAT TTT GAT TAT CTC	388
	Ser Val Ser Leu Thr Ile Ala Glu Ser Ser Met Asp Phe Asp Tyr Leu	
5	105 110 115	
	GCC GGA GAT CAT CAG AGG GAT TCT TAT AAA TTC AAC GAT TTG ATT CCG	436
	Ala Gly Asp His Gln Arg Asp Ser Tyr Lys Phe Asn Asp Leu Ile Pro	
10	120 125 130	
	CAG CTG CCA GAA CCG ATT GTA ACC TCC GGC GAC GAA GTA TTA CCA CTT	484
	Gln Leu Pro Glu Pro Ile Val Thr Ser Gly Asp Glu Val Leu Pro Leu	
	135 140 145 150	
15	TTT GCT TTA CAG GTG ACG GTG TTC TCC AAC ACC GGT ATA TGC ATT GGA	532
	Phe Ala Leu Gln Val Thr Val Phe Ser Asn Thr Gly Ile Cys Ile Gly	
	155 160 165	
20	CGC AAT CTT CAT CAA GTT CTT GGT GAT GCC AGT TCT TTT CTG CAT TTT	580
	Arg Asn Leu His Gln Val Leu Gly Asp Ala Ser Ser Phe Leu His Phe	
	170 175 180	
	AAT AAA TTA TGG GTT TTG GTT GAC AAA TCC AAT GGA GAT TCA TTA AAG	628
25	Asn Lys Leu Trp Val Leu Val Asp Lys Ser Asn Gly Asp Ser Leu Lys	
	185 190 195	
	TTC CTT CCA CTT TCT TCT CTA CCT ATG TAC GAC AGA TCT GTG GTG CAA	676
30	Phe Leu Pro Leu Ser Ser Leu Pro Met Tyr Asp Arg Ser Val Val Gln	
	200 205 210	
	GAT CCA TTT CAT ATT CGT CGA AAA ATC TAC AAT GAA AGA AAA CTG CTC	724
	Asp Pro Phe His Ile Arg Arg Lys Leu Tyr Asn Glu Arg Lys Leu Leu	
35	215 220 225 230	
	AAA TCT CAG GGC ACA CCT ACT GTT CTA AAT CCA GCA ATT TCT AAA GAT	772
	Lys Ser Gln Gly Thr Pro Thr Val Leu Asn Pro Ala Ile Ser Lys Asp	
	235 240 245	
40	GAA GTT CGA GCC ACC TTC ATC CTA CAC CCT ATT GAT ATC ATG AAG CTC	820
	Glu Val Arg Ala Thr Phe Ile Leu His Pro Ile Asp Ile Met Lys Leu	
	250 255 260	
45	AAG AAA TTC ATT TCG TCA AAA AAT CGC AAC TTA ACC GGT AGT AGT AAT	868
	Lys Lys Phe Ile Ser Ser Lys Asn Arg Asn Leu Thr Gly Ser Ser Asn	
	265 270 275	
	TAT AAT CTG TCA ACT TTC ACG GTG ACA TCT GCA CTG ATC TGG ACA TGC	916
50	Tyr Asn Leu Ser Thr Phe Thr Val Thr Ser Ala Leu Ile Trp Thr Cys	
	280 285 290	

55



**EP 0 810 287 A1**

	TTG TCG AAA TCA TTA GAC ACC GTC GTA AGA GAG AAG GTC GAA GAG GAT	964
5	Leu Ser Lys Ser Leu Asp Thr Val Val Arg Glu Lys Val Glu Glu Asp 295 300 305 310	
	AAA CAT GCA GCA AAC TTA TGT GCT TTC ATC AAC TGC CGA CAA CGT TTT	1012
	Lys His Ala Ala Asn Leu Cys Ala Phe Ile Asn Cys Arg Gln Arg Phe 315 320 325	
10	GCT CCG CCG ATA CCT CAA AAT TAC TTT GGA AAT TGC ATA GTG CCT TGT Ala Pro Pro Ile Pro Gln Asn Tyr Phe Gly Asn Cys Ile Val Pro Cys 330 335 340	1060
15	ATG GTG GGA TCG ACT CAT GAG CAA CTT GTA GGA AAT GAA GGG TTG TCG Met Val Gly Ser Thr His Glu Gln Leu Val Gly Asn Glu Gly Leu Ser 345 350 355	1108
20	GTA GCT GCA ACC GCC ATC GGA GAT GCT ATC CAT AAG AGG TTA CAT GAC Val Ala Ala Thr Ala Ile Gly Asp Ala Ile His Lys Arg Leu His Asp 360 365 370	1156
	TAC GAA GGA ATT CTG AGA GGA GAT TGG ATA TCG CCG CCC CGA TCA ACA Tyr Glu Gly Ile Leu Arg Gly Asp Trp Ile Ser Pro Pro Arg Ser Thr 375 380 385 390	1204
25	TCT GCG GCA CCA AGG TCG ACG CTC ATT TAT GTC GTT GGA TCC GCA CAA Ser Ala Ala Pro Arg Ser Thr Leu Ile Tyr Val Val Gly Ser Ala Gln 395 400 405	1252
30	CGC AAT GTG CAT GAT TTT GAT GCA GAT TTT GGT TGG GGA AAG CTT GAA Arg Asn Val His Asp Phe Asp Ala Asp Phe Gly Trp Gly Lys Leu Glu 410 415 420	1300
35	AAG CAT GAA TCT GTT TCA ACT AAT CCT TCG GCA ACA CTA ATT TTG ATC Lys His Glu Ser Val Ser Thr Asp Pro Ser Ala Thr Leu Ile Leu Ile 425 430 435	1348
40	TCT CGG TCC AGA AGA TTT AAA GGA GCA CTT GAG CTT GGC ATT TCT TTG Ser Arg Ser Arg Arg Phe Lys Gly Ala Leu Glu Leu Gly Ile Ser Leu 440 445 450	1396
	CCT AAG AAT AGG ATG GAC GCA TTT GCC ACC ATT TTT ACG AAT TTC ATC Pro Lys Asn Arg Met Asp Ala Phe Ala Thr Ile Phe Thr Asn Phe Ile 455 460 465 470	1444
45	AAT AGT CTC CAT GTG AGG AGC CCT TTG TAAGAAAAAA GTGGTATCAA Asn Ser Leu His Val Arg Ser Pro Leu 475 479	1491
50	TGTATAAAAA AGACAGACAA GTTATGATGC AACAAATGTT TTAGGACATT ACAAATCCAT	1551

EP 0 810 287 A1

GGGAAGATGT ATCAAACTCA TCTCTCTATA TATATATATT CAATTGTTTT AAAAAAAAAA 1611  
AAAAAAAAAA A 1622

SEQ ID NO: 3

SEQUENCE LENGTH: 1605

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL SEQUENCE: No

ANTI-SENSE: No

ORIGINAL SOURCE

ORGANISM: Petunia hybrida

TISSUE TYPE: petal

IMMEDIATE SOURCE

LIBRARY: cDNA library

CLONE: pPAT48

SEQUENCE DESCRIPTION

IGTCGACGAA ATCCATTICA TTTCCTCTTC TTTCCTGTTT TTCTAATTC GTCATCATTG 60

TTATCC ATG GCA GGT GAA GTA GCA AAA CAA GAA GTT ACA AAA GTG AAA 108

Met Ala Gly Glu Val Ala Lys Gln Glu Val Thr Lys Val Lys

1 5 10

GTC CTG AAA AAA ACA AAC GTG AAA CCA CAT AAA CCA CTA GGA AAA AAA 156

Val Leu Lys Lys Thr Asn Val Lys Pro His Lys Pro Leu Gly Lys Lys

15 20 25 30

GAG TGT CAA TTG GTA ACA TTT GAT CTT CCT TAC CTA GCT TTC TAT TAC 204

Glu Cys Gln Leu Val Thr Phe Asp Leu Pro Tyr Leu Ala Phe Tyr Tyr

35 40 45

AAC CAA AAA TTT CTC ATC TAT AAA GGT GCT GAA AAC TTT GAC GAG ACG 252

Asn Gln Lys Phe Leu Ile Tyr Lys Gly Ala Glu Asn Phe Asp Glu Thr

50 55 60

GTG GAA AAA ATT AAA GAT GGA CTG GCC TTA GTA TTG GTG GAT TTC TAT 300

Val Glu Lys Ile Lys Asp Gly Leu Ala Leu Val Leu Val Asp Phe Tyr

65 70 75

CAA CTA GCT GGG AAA CTT GGA AAA GAT GAA GAA GGG GTT TTC AGG GTG 348

Gln Leu Ala Gly Lys Leu Gly Lys Asp Glu Glu Gly Val Phe Arg Val

80 85 90

EP 0 810 287 A1

	GAA TAC GAC GAT GAC ATG GAT GGT GTA GAG CTG ACA GTG GCT GTT GCA	396
5	Glu Tyr Asp Asp Asp Met Asp Gly Val Glu Val Thr Val Ala Val Ala	
	95 100 105 110	
	GAA GAG ATA GAA GTT GCA GAT CTT ACT GAT GAA GAA GGC ACC ACC AAA	444
	Glu Glu Ile Glu Val Ala Asp Leu Thr Asp Glu Glu Gly Thr Thr Lys	
10	115 120 125	
	TTC CAG GAC TTG ATT CCT TGT AAT AAA ATC TTG AAT TTG GAA GGG CTT	492
	Phe Gln Asp Leu Ile Pro Cys Asn Lys Ile Leu Asn Leu Glu Gly Leu	
	130 135 140	
15	CAT CGC CCT CTT CTA GCT GTG CAG CTC ACC AAG CTC AAG GAC GGG CTC	540
	His Arg Pro Leu Leu Ala Val Gln Leu Thr Lys Leu Lys Asp Gly Leu	
	145 150 155	
20	ACC ATG GGA TTA GCA TTT AAC CAT GCT GTG CTG GAT GGT ACT TCG ACG	588
	Thr Met Gly Leu Ala Phe Asn His Ala Val Leu Asp Gly Thr Ser Thr	
	160 165 170	
	TGG CAC TTT ATG ACC TCG TGG TCC GAG CTT TGC TGT GGG TCC ACC TCA	636
25	Trp His Phe Met Thr Ser Trp Ser Glu Leu Cys Cys Gly Ser Thr Ser	
	175 180 185 190	
	ATT TCT GTC CCA CCA TTC CTT GAA CGA ACC AAG GCT CGT AAC ACT CGA	684
	Ile Ser Val Pro Pro Phe Leu Glu Arg Thr Lys Ala Arg Asn Thr Arg	
30	195 200 205	
	GTC AAG CTC AAC CTC TCT CAA CCA TCA GAT GCA CCC GAA CAT GCT AAG	732
	Val Lys Leu Asn Leu Ser Gln Pro Ser Asp Ala Pro Glu His Ala Lys	
35	210 215 220	
	TCA GCA ACC AAC GGT GAT GTC CCG GCC AAC GTA GAC CCA CCT CTT CGC	780
	Ser Ala Thr Asn Gly Asp Val Pro Ala Asn Val Asp Pro Pro Leu Arg	
	225 230 235	
40	GAA AGA GTA TTC AAG TTC TCG GAG TTA GCA ATT GAC AAA ATC AAG TCA	828
	Glu Arg Val Phe Lys Phe Ser Glu Leu Ala Ile Asp Lys Ile Lys Ser	
	240 245 250	
45	ACA GTC AAT GCC AAC TCA GGA GAG ACG CCA TTC TCC ACA TTC CAA TCA	876
	Thr Val Asn Ala Asn Ser Gly Glu Thr Pro Phe Ser Thr Phe Gln Ser	
	255 260 265 270	
	CTC TCC GCA CAC GTG TGG CTA GCC GTC ACA CGT GCG GCG CAA CTC AAG	924
50	Leu Ser Ala His Val Trp Leu Ala Val Thr Arg Ala Arg Gln Leu Lys	
	275 280 285	

EP 0 810 287 A1

CCC GAG GAC TAC ACT GTG TAC ACT GTG TTT GCT GAT TGC AGG AAA AGG 972  
 Pro Glu Asp Tyr Thr Val Tyr Thr Val Phe Ala Asp Cys Arg Lys Arg  
 290 295 300

5 GTT GAT CCT CCA ATG CCA GAA AGT TAC TTC GGC AAC CTA ATT CAG GCA 1020  
 Val Asp Pro Pro Met Pro Glu Ser Tyr Phe Gly Asn Leu Ile Gln Ala  
 305 310 315

10 ATT TTC ACA GTG ACC GCG CCA GGT TTG TTA CTA GCA AGC CCG ATC GAG 1068  
 Ile Phe Thr Val Thr Ala Ala Gly Leu Leu Ala Ser Pro Ile Glu  
 320 325 330

15 TTC GCT GGT GGG ATG ATA CAA CAA GCG ATC GTG AAG CAT GAC GCT AAG 1116  
 Phe Ala Gly Gly Met Ile Gln Gln Ala Ile Val Lys His Asp Ala Lys  
 335 340 345 350

20 CCC ATT GAT GAA AGA AAC AAG GAG TGG GAG AGC AAC CCG AAG ATC TTT 1164  
 Ala Ile Asp Glu Arg Asn Lys Glu Trp Glu Ser Asn Pro Lys Ile Phe  
 355 360 365

25 CAG TAC AAA GAT GCT GGA GTG AAC TGT GTT GCT GTT GGA AGT TCG CCA 1212  
 Gln Tyr Lys Asp Ala Gly Val Asn Cys Val Ala Val Gly Ser Ser Pro  
 370 375 380

AGG TTC AAG GTT TAC GAC GTG GAT TTT GGA TGG GGA AAG CCA GAG AGT 1260  
 Arg Phe Lys Val Tyr Asp Val Asp Phe Gly Trp Gly Lys Pro Glu Ser  
 385 390 395

30 GTG AGG AGT GGT TCG AAC AAT AGG TTT GAT GGA ATG GTC TAT TTG TAC 1308  
 Val Arg Ser Gly Ser Asn Asn Arg Phe Asp Gly Met Val Tyr Leu Tyr  
 400 405 410

35 CAA GGC AAA AAT GGA GGA AGA AGC ATT GAT GTG GAG ATT AGT TTG GAA 1356  
 Gln Gly Lys Asn Gly Gly Arg Ser Ile Asp Val Glu Ile Ser Leu Glu  
 415 420 425 430

40 GCA AAT GCT ATC CAG AGG TTG GAG AAA GAT AAA GAG TTC CTC ATG GAA 1404  
 Ala Asn Ala Met Glu Arg Leu Glu Lys Asp Lys Glu Phe Leu Met Glu  
 435 440 445

ACT GCT TAATTTGCTT AGCTTGGACT CAACTGGCTA CACTTTATTT ATGAGCTGCT 1460  
 Thr Ala

45 ATGACTCACA TGCATGTATG TTTATTTTTT TTGGAGGGGT TCITTCCTTT TATTGTTTTC 1520  
 TATGTTTTTT CTITCTTGTA CGTTATGAAG AGAAACCGAG TATAAAGGAA TAATGTTTTC 1580  
 AGTTATTAAA AAAAAAAAAA AAAA 1605

50 SEQ ID NO: 4  
 SEQUENCE LENGTH: 1479

SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: double  
 5 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL SEQUENCE: No  
 ANTI-SENSE: No  
 10 ORIGINAL SOURCE  
 ORGANISM: *Perilla ocimoides*  
 TISSUE TYPE: leaf  
 15 IMMEDIATE SOURCE  
 LIBRARY: cDNA library  
 CLONE: pSAT208  
 SEQUENCE DESCRIPTION  
 20 CC GTG ATC GAA ACG TGT AGA GTT GGG CCG CCG CCG GAC TCG GTG GCG 47  
 Val Ile Glu Thr Cys Arg Val Gly Pro Pro Pro Asp Ser Val Ala  
 1 5 10 15  
 25 GAG CAA TCG GTG CCG CTC ACA TTC TTC GAC ATG ACG TGG CTG CAT TTT 95  
 Glu Gln Ser Val Pro Leu Thr Phe Phe Asp Met Thr Trp Leu His Phe  
 20 25 30  
 30 CAT CCC ATG CTT CAG CTC CTC TTC TAC GAA TTC CCT TGT TCC AAG CAA 143  
 His Pro Met Leu Gln Leu Leu Phe Tyr Glu Phe Pro Cys Ser Lys Gln  
 35 40 45  
 35 CAT TTT TCA GAA TCC ATC GTT CCA AAA CTC AAA CAA TCT CTC TCT AAA 191  
 His Phe Ser Glu Ser Ile Val Pro Lys Leu Lys Gln Ser Leu Ser Lys  
 50 55 60  
 40 ACT CTC ATA CAC TTC TTC CCT CTC TCA TGC AAT TTA ATC TAC CCT TCA 239  
 Thr Leu Ile His Phe Phe Pro Leu Ser Cys Asn Leu Ile Tyr Pro Ser  
 65 70 75  
 45 TCC CCG GAG AAA ATG CCG GAG TTT CGG TAT CTA TCC GGG GAC TCG GTT 287  
 Ser Pro Glu Lys Met Pro Glu Phe Arg Tyr Leu Ser Gly Asp Ser Val  
 80 85 90 95  
 50 TCT TTC ACC ATC GCA GAA TCT AGC GAC GAC TTC GAT GAT CTC GTC GGA 335  
 Ser Phe Thr Ile Ala Glu Ser Ser Asp Asp Phe Asp Asp Leu Val Gly  
 100 105 110  
 55 AAT CGT CCA GAA TCT CCC GTT AGG CTC TAC AAC TTT GTC CCT AAA TTG 383  
 Asn Arg Pro Glu Ser Pro Val Arg Leu Tyr Asn Phe Val Pro Lys Leu  
 115 120 125

55

EP 0 810 287 A1

CCG CCC ATT GTC GAA GAA TCC GAT AGA AAA CTC TTC CAA GTT TTC GCC 431  
 Pro Pro Ile Val Glu Glu Ser Asp Arg Lys Leu Phe Gln Val Phe Ala  
 5 130 135 140  
 GTG CAG GTG ACT CTT TTC CCA GGC CGA GGC GTC GGT ATT GGA ATA GCA 479  
 Val Gln Val Thr Leu Phe Pro Gly Arg Gly Val Gly Ile Gly Ile Ala  
 10 145 150 155  
 ACG CAT CAC ACC GTT AGC GAC GCC CCG TCG TTT CTC GCG TTT ATA ACG 527  
 Thr His His Thr Val Ser Asp Ala Pro Ser Phe Leu Ala Phe Ile Thr  
 160 165 170 175  
 GCT TGG TCT TCA ATG AGC AAA CAC ATT GAA AAT GAA GAT GAA GAT GAA 575  
 Ala Trp Ser Ser Met Ser Lys His Ile Glu Asn Glu Asp Glu Asp Glu  
 180 185 190  
 GAA TTT AAA TCT TTG CCA GTT TTC GAT AGA TCC GTC ATA AAA TAT CCG 623  
 Glu Phe Lys Ser Leu Pro Val Phe Asp Arg Ser Val Ile Lys Tyr Pro  
 195 200 205  
 ACG AAA TTT GAC TCC ATT TAT TGG AGA AAC GCG CTA AAA TTT CCT TTG 671  
 Thr Lys Phe Asp Ser Ile Tyr Trp Arg Asn Ala Leu Lys Phe Pro Leu  
 210 215 220  
 CAA TCT CGT CAT CCC TCA TTA CCG ACG GAC CGC ATT CGA ACC ACG TTC 719  
 Gln Ser Arg His Pro Ser Leu Pro Thr Asp Arg Ile Arg Thr Thr Phe  
 225 230 235  
 GTT TTC ACC CAA TCC AAA ATT AAG AAA TTG AAG GGT TGG ATT CAG TCC 767  
 Val Phe Thr Gln Ser Lys Ile Lys Lys Leu Lys Gly Trp Ile Gln Ser  
 240 245 250 255  
 AGA GTT CCA ACT TTA GTC CAT CTC TCA TCT TTT GTA GCG ATT GCA GCT 815  
 Arg Val Pro Ser Leu Val His Leu Ser Ser Phe Val Ala Ile Ala Ala  
 260 265 270  
 TAT ATG TGG GGT GGC ATA ACG AAA TCA TTC ACA GCA GAT GAA GAC CAA 863  
 Tyr Met Trp Ala Gly Ile Thr Lys Ser Phe Thr Ala Asp Glu Asp Gln  
 275 280 285  
 GAC AAC GAG GAT GCA TTT TTC TTG ATT CCG GTC GAT CTA AGG CCA CGA 911  
 Asp Asn Glu Asp Ala Phe Phe Leu Ile Pro Val Asp Leu Arg Pro Arg  
 290 295 300  
 TTA GAT CCG CCG GTT CCT GAA AAT TAC TTC GGG AAC TGC TTA TCG TAC 959  
 Leu Asp Pro Pro Val Pro Glu Asn Tyr Phe Gly Asn Cys Leu Ser Tyr  
 305 310 315

55

GCG CTG CCG AGA ATG CGG CGG CGA GAG CTG GTG GGA GAG AAA GGG GTG 1007  
 Ala Leu Pro Arg Met Arg Arg Arg Glu Leu Val Gly Glu Lys Gly Val  
 5 320 325 330 335  
 TTT CTG GCA GCT GAG GTA ATC CCG CGC GAG ATA AAA AAA AGG ATC AAC 1055  
 Phe Leu Ala Ala Glu Val Ile Ala Ala Glu Ile Lys Lys Arg Ile Asn  
 340 345 350  
 10 GAC AAG AGA ATA TTA GAA ACG CTG GAC AAA TGG TCG CCG GAC ATT CGT 1103  
 Asp Lys Arg Ile Leu Glu Thr Val Glu Lys Trp Ser Pro Glu Ile Arg  
 355 360 365  
 15 AAA CGC TTG CAG AAA TCA TAT TTT TCG GTG GCA GCA TCG AGC AAG CTA 1151  
 Lys Ala Leu Gln Lys Ser Tyr Phe Ser Val Ala Gly Ser Ser Lys Leu  
 370 375 380  
 20 GAT CTT TAC GGT GCA GAT TTT GGA TGG GGG AAG GCG AGA AAG GAA GAA 1199  
 Asp Leu Tyr Gly Ala Asp Phe Gly Trp Gly Lys Ala Arg Lys Gln Glu  
 385 390 395  
 ATA TTG TCG ATT GAT GGC GAG AAA TAT GCA ATG ACR CTT TCT AAA GCC 1247  
 Ile Leu Ser Ile Asp Gly Glu Lys Tyr Ala Met Thr Leu Cys Lys Ala  
 25 400 405 410 415  
 AGG GAT TTC GAA GGA GGA TTG GAG GTT TGC TTG TCT TTG CCT AAG GAC 1295  
 Arg Asp Phe Glu Gly Gly Leu Glu Val Cys Leu Ser Leu Pro Lys Asp  
 420 425 430  
 30 AAA ATC GAT CCT TTT GCT GCT TAT TTT TCA CTG GGA ATT AAT GGT 1340  
 Lys Met Asp Ala Phe Ala Ala Tyr Phe Ser Leu Gly Ile Asn Gly  
 435 440 446  
 35 TAATAAATGT ATGTAATTAA ACTAATATTA TTATGTAACA ATTAATTAAG TCGTTGAGTAA 1400  
 CGTGAAGAAT AATCCCTATT ATATATTTAT GATTGGTTC AAATAAAGTG TAAAGCCTCT 1460  
 TCAAAAAAAAA AAAAAAAAAA 1479  
 SEQ ID NO: 5  
 40 SEQUENCE LENGTH: 1508  
 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: double  
 45 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL SEQUENCE: No  
 ANTI-SENSE: No  
 50 ORIGINAL SOURCE  
 ORGANISM: Senecio cruentus

TISSUE TYPE: petal

IMMEDIATE SOURCE

LIBRARY: cDNA library

CLONE: pCAT8

## SEQUENCE DESCRIPTION

```

5      TG AAC ATT CTC GAA CAT GCC CGA ATA TCG GCC CCC TCG GGC ACC ATC      47
10      Asn Ile Leu Glu His Ala Arg Ile Ser Ala Pro Ser Gly Thr Ile
      1           5           10          15
      GGC CAT CGC TCG TTA TCT CTT ACT TTC TTC GAC ATT ACT TGG CTA CTC      95
15      Gly His Arg Ser Leu Ser Leu Thr Phe Phe Asp Ile Thr Trp Leu Leu
      20           25           30
      TTC CCT CCG GTC CAC CAT CTT TTC TTC TAT GAC TTT CCA CAT TCT AAA      143
20      Phe Pro Pro Val His His Leu Phe Phe Tyr Asp Phe Pro His Ser Lys
      35           40           45
      TCC CAT TTC ATG GAC ACT ATT GTT CCC AGG CTA AAA CAA TCT TTA TCG      191
25      Ser His Phe Met Asp Thr Ile Val Pro Arg Leu Lys Gln Ser Leu Ser
      50           55           60
      GTC ACT CTT CAA CAT TTT TTC CCG TTT GCT AGT AAT TTG ATT GTA TTT      239
30      Val Thr Leu Gln His Phe Phe Pro Phe Ala Ser Asn Leu Ile Val Phe
      65           70           75
      CCT AAC ACT GAT GGT TCG GGT TTT AAT AAA AAA CCA GAA ATA AAA CAC      287
      Pro Asn Thr Asp Gly Ser Gly Phe Asn Lys Lys Pro Glu Ile Lys His
      80           85           90           95
35      GTT GAA GGT GAT TCT GTT GTG GTT ACT TTT GCA GAA TGT TGT CTT GAC      335
      Val Glu Gly Asp Ser Val Val Val Thr Phe Ala Glu Cys Cys Leu Asp
      100          105          110
40      TTT AAT AAT TTG ACA GGA AAT CAT CCT CGA AAA TGT GAA AAC TTT TAT      383
      Phe Asn Asn Leu Thr Gly Asn His Pro Arg Lys Cys Glu Asn Phe Tyr
      115          120          125
45      CCA CTT GTA CCT TCA TTG GGA AAT GCA ATC AAA TTA TGT GAT TGC GTC      431
      Pro Leu Val Pro Ser Leu Gly Asn Ala Ile Lys Leu Cys Asp Cys Val
      130          135          140
50      ACG GTC CCA CTT TTT TCA CTT CAA GTG ACG TTT TTT CCG GGC TCG GGT      479
      Thr Val Pro Leu Phe Ser Leu Gln Val Thr Phe Phe Pro Gly Ser Gly
      145          150          155

```



EP 0 810 287 A1

5      ATA TCA CTA GGA ATG ACG AAT CAT CAT AGC CTT GGT GAC GGT AGC ACG      527  
       Ile Ser Leu Gly Met Thr Asn His His Ser Leu Gly Asp Ala Ser Thr  
       160                      165                      170                      175  
       CGG TTC AAC TTT TTG AAA GGG TGG ACT TCG ATT ATT CAA TCT GGT GTA      575  
       Arg Phe Asn Phe Leu Lys Gly Trp Thr Ser Ile Ile Gln Ser Gly Val  
       180                      185                      190  
 10      GAT CGG TCT TTT TTA ACG AAA GGA TCT CCA CCG GTT TTT GAT AGA TTG      623  
       Asp Arg Ser Phe Leu Thr Lys Gly Ser Pro Pro Val Phe Asp Arg Leu  
       195                      200                      205  
 15      ATT AAC ATC CCA CAT TTA GAT GAA AAT AAG TTG AGA CAT ACA AGG CTC      671  
       Ile Asn Ile Pro His Leu Asp Glu Asn Lys Leu Arg His Thr Arg Leu  
       210                      215                      220  
 20      GAA AGT TTT TAT AAA CCT TCG AGC CTT GTT GGT CCC ACT GAT AAA GTT      719  
       Glu Ser Phe Tyr Lys Pro Ser Ser Leu Val Gly Pro Thr Asp Lys Val  
       225                      230                      235  
       CGG TCA ACG TTT GTG TTG ACC CGA ACT AAT ATC AAT CTA CTA AAG AAA      767  
 25      Arg Ser Thr Phe Val Leu Thr Arg Thr Asn Ile Asn Leu Leu Lys Lys  
       240                      245                      250                      255  
       AAG GTC TTA ACC CAA GTG CCA AAC TTG GAG TAC ATG TCA TCT TTT ACG      815  
 30      Lys Val Leu Thr Gln Val Pro Asn Leu Glu Tyr Met Ser Ser Phe Thr  
       260                      265                      270  
       GTA ACT TGT GGT TAT ATA TGG AGT TGC ATA GCG AAA TCA CTC GTA AAA      863  
       Val Thr Cys Gly Tyr Ile Trp Ser Cys Ile Ala Lys Ser Leu Val Lys  
       275                      280                      285  
 35      ATA GGA GAA AGA AAG GGC GAA GAC GAG TTA GAA CAG TTC ATA ATC ACC      911  
       Ile Gly Glu Arg Lys Gly Glu Asp Glu Leu Glu Gln Phe Ile Ile Thr  
       290                      295                      300  
 40      ATT GAT TGT CGA TCT CGT CTT GAT CCA CCA ATT CCC ACA GCC TAC TTT      959  
       Ile Asp Cys Arg Ser Arg Leu Asp Pro Pro Ile Pro Thr Ala Tyr Phe  
       305                      310                      315  
 45      GGT AAC TGT GGT GCA CCA TGT GTC CCG ACC TTA AAA AAT GTC GTT TTG      1007  
       Gly Asn Cys Gly Ala Pro Cys Val Pro Thr Leu Lys Asn Val Val Leu  
       320                      325                      330                      335  
 50      ACT AGC GAA AAT GGG TAT GCA CTT GGT GCT AAA GTA ATT GGA GAG TCT      1055  
       Thr Ser Glu Asn Gly Tyr Ala Leu Gly Ala Lys Val Ile Gly Glu Ser  
       340                      345                      350

55

EP0810287A1

ATA TGC AAA ATG ATA TAT AAT AAG GAC GGA ATC TTG AAA GAT GCC GCG 1103  
Ile Cys Lys Met Ile Tyr Asn Lys Asp Gly Ile Leu Lys Asp Ala Ala

355 360 365

AGA TGG CAT GAA CCT TTC ATG ATC CCG GCT AGG AAG ATT GGT GTT GCT 1151  
Arg Trp His Glu Pro Phe Met Ile Pro Ala Arg Lys Ile Gly Val Ala

370 375 380

GGT ACA CCT AAG CTC AAC TTG TAC GAC TTT GAT TTT GGG TGG GGG AAG 1199  
Gly Thr Pro Lys Leu Asn Leu Tyr Asp Phe Asp Phe Gly Trp Gly Lys

385 390 395

CCG ATA AAG TAT GAG ACT GTT TCA ATA GAC TAT AAT ACG TCG ATT TCT 1247  
Pro Ile Lys Tyr Glu Thr Val Ser Ile Asp Tyr Asn Thr Ser Ile Ser

400 405 410 415

ATA AAT GCA AGC AAA ACA TCA GCA CAA GAT CTT GAA ATT GGA TTG AGT 1295  
Ile Asn Ala Ser Lys Thr Ser Ala Gln Asp Leu Glu Ile Gly Leu Ser

420 425 430

CTA CCG AGT ATG CAA ATG GAG GCG TTT TCT AGC ATC TTT GAT GAA GGA 1343  
Leu Pro Ser Met Gln Met Glu Ala Phe Ser Ser Ile Phe Asp Glu Gly

435 440 445

TTA GAG AGT CAA GTT TCA TTG TAGATCATCG TCCCTTTTT GTGTGCATCA 1394  
Leu Glu Ser Gln Val Ser Leu

450 454

AGTTTCGTGTC GTTTTATGA GTTGCCACTG TTCTATTCTT TAAGTATACC TTTCGACTAT 1454  
GTTTGAAGA TGCAACGATA TAAATGAAA AAAAAAAAAA AAAAAAAAAA AAAA 1508

SEQ ID NO: 6

SEQUENCE LENGTH: 1521

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL SEQUENCE: No

ANTI-SENSE: No

ORIGINAL SOURCE

ORGANISM: *Lavandula angustifolia*

TISSUE TYPE: petal

IMMEDIATE SOURCE

LIBRARY: cDNA library

CLONE: pLAT21

EP 0 810 287 A1

SEQUENCE DESCRIPTION	
-TG ACC ACC CTC CTC GAA TCC TCC CGA GTG GCG CCG CCT CCA GGC ACG	47
Xxx Thr Thr Leu Leu Glu Ser Ser Arg Val Ala Pro Pro Pro Gly Thr	
1 5 10 15	
GTG GCT GAG CAG TCA CTC CCG CTC ACC TTC TTC GAC ATG ACG TGG CTG	95
Val Ala Glu Gln Ser Leu Pro Leu Thr Phe Phe Asp Met Thr Trp Leu	
20 25 30	
CAT TTC CAC CCC ATG CTT CAG CTT CTC TTC TAC GAA CTC GCC TGT TCC	143
His Phe His Pro Met Leu Gln Leu Leu Phe Tyr Glu Leu Pro Cys Ser	
35 40 45	
AAA CCC GCC TTC CTC GAA ACC GTC GTT CCG AAA CTC AAA CAA TCC TTA	191
Lys Pro Ala Phe Leu Glu Thr Val Val Pro Lys Leu Lys Gln Ser Leu	
50 55 60	
TCT CTA ACC CTC AAA CAC TTC TTC CCC CTT TCA TGC AAT CTA ATC TAC	239
Ser Leu Thr Leu Lys His Phe Phe Pro Leu Ser Cys Asn Leu Ile Tyr	
65 70 75	
CCT CTA TCG CCG GAG AAA ATG CCG GAG TTC CCG TAT CAG AAC GGT GAC	287
Pro Leu Ser Pro Glu Lys Met Pro Glu Phe Ser Val Ser Phe Thr Ile	
80 85 90 95	
TCG GTT TCT TTC ACG ATT ATG GAG TCT GTC GGA GAT CAT CCG CAT TCC	335
Met Glu Ser Ser Asp Asp Tyr Glu Asp Val Gly Asp His Pro His Ser	
100 105 110	
GCT CAT AAA TAC TAC TGC TTT GCC CCT AGC GAC GAT TAT GAA GAT CTC	383
Ala His Lys Tyr Tyr Cys Phe Ala Gln Leu Pro Pro Ile Val Glu Glu	
115 120 125	
CAG CTC CCG CCG ATA GTC GAG GAA TCT GAT CCG AAA TTG TTT CAA GTT	431
Ser Asp Arg Lys Leu Phe Gln Val Pro Leu Arg Tyr Gln Asn Gly Asp	
130 135 140	
TTA GCC GTG CAA GTG ACT CTG TTT CCC GGT CCG GGG GTG TGC ATC GGA	479
Leu Ala Val Gln Val Thr Leu Phe Pro Gly Arg Gly Val Cys Ile Gly	
145 150 155	
ATA ACG ACG CAC CAC ACC GTT AGC GAT GCT CCA TCG TTT GTA GGG TTT	527
Ile Thr Thr His His Thr Val Ser Asp Ala Pro Ser Phe Val Gly Phe	
160 165 170 175	
ATG AAG AGT TGG GCT TCC ATC ACT AAA TTC GGA GGA GAT GAT GAA TTC	575
Met Lys Ser Trp Ala Ser Ile Thr Lys Phe Gly Gly Asp Asp Glu Phe	
180 185 190	

EP 0 810 287 A1

	TTG GAC GGA AAA GGT GAA TGT TTG CCG GTT TTC GAC CGA TCG CTC GTG	623
5	Leu Asp Gly Lys Gly Glu Cys Leu Pro Val Phe Asp Arg Ser Leu Val	
	195 200 205	
	AAT TAT CCG CCT AAA TTG GAC ACA TAT TTA TGG AAC AAC GCG CAG AAA	671
	Asn Tyr Pro Pro Lys Leu Asp Thr Tyr Leu Trp Asn Ala Gln Lys	
10	210 215 220	
	CGT CCG TTG GAA TCG CAG CAT CCA TCT TTA CCG ACG GAT CGG ATT CGA	719
	Arg Pro Leu Glu Ser Gln His Pro Ser Leu Pro Thr Asp Arg Ile Arg	
	225 230 235	
15	GCT ACC TAC CTT TTC ACC CAA TCT GAA ATT AAG AAA TTG AAG GGT TTG	767
	Ala Thr Tyr Leu Phe Thr Gln Ser Glu Ile Lys Lys Leu Lys Gly Leu	
	240 245 250 255	
20	ATT CAG AGA AAA GCC CCA AAT GTA GTT AAT CTC TCT TCC TTC GTC GCG	815
	Ile Gln Arg Lys Ala Pro Asn Val Val Asn Leu Ser Ser Phe Val Ala	
	260 265 270	
	ATC GCA GCT TAT ATC TGG ACC GGC ATC GCC AAA TCG GTC GGA GAT TAC	863
25	Ile Ala Ala Tyr Ile Trp Thr Gly Ile Ala Lys Ser Val Gly Asp Tyr	
	275 280 285	
	AAA GAC GTG GAT GAC GAC AAA GCG GCT TTC TTT TTA ATT CCG ATC GAT	911
30	Lys Asp Val Asp Asp Asp Lys Arg Ala Phe Phe Leu Ile Pro Ile Asp	
	290 295 300	
	TTA AGG CCG CGT TTG GAT CCG CCG GCT CCG GGG AAC TAC TTC GGA AAC	959
	Leu Arg Pro Arg Leu Asp Pro Pro Ala Pro Gly Asn Tyr Phe Gly Asn	
35	305 310 315	
	TGT CTA TCG TTT GCG ATG GCG AAG ATC CTG CCG CCG GAT TTC GTC GGA	1007
	Cys Leu Ser Phe Ala Met Ala Lys Ile Leu Arg Arg Asp Leu Val Gly	
	320 325 330 335	
40	GAT GAA GGG GTG TTT CCG GCA GCT GAG GCG ATC GCG CCG GAA ATA GAG	1055
	Asp Glu Gly Val Phe Arg Ala Ala Glu Ala Ile Ala Ala Glu Ile Glu	
	340 345 350	
45	AAG AGC ACG AGC GAC AAG AAG ATT CTA GAA ACT GTG GAC AAC TGG CCG	1103
	Lys Arg Thr Ser Asp Lys Lys Ile Leu Glu Thr Val Glu Asn Trp Pro	
	355 360 365	
50	TCT GAG ATT CCG GAA GCC TTG CAA AAC TGT TAT TTC TCG GTG GCG GGA	1151
	Ser Glu Ile Arg Glu Ala Leu Gln Asn Cys Tyr Phe Ser Val Ala Gly	
	370 375 380	

EP 0 810 287 A1

TCG AGC AGG CTT GAT CTT TAC GGC GCG GAT TTT GGA TGG GGT AAG GCG 1199  
 Ser Ser Arg Leu Asp Leu Tyr Gly Ala Asp Phe Gly Trp Gly Lys Ala  
 5 385 390 395

GTG AAG CAA GAG ATA CTG TCG ATT GAT GGA GAG AAG TTT ACG ATC TCG 1247  
 Val Lys Gln Glu Ile Leu Ser Ile Asp Gly Glu Lys Phe Thr Met Ser  
 10 400 405 410 415

TTG TGT AAA CCG AGG GAT GCT GCC GGA GGA TTG GAG GTT GGA TTG TCT 1295  
 Leu Cys Lys Pro Arg Asp Ala Ala Gly Gly Leu Glu Val Gly Leu Ser  
 420 425 430

TTG CCA AAG GAG GAA TTG CAA GCT TTT GAT TAT TTT CCG GAG GGA 1343  
 Leu Pro Lys Glu Glu Leu Gln Ala Phe Asp Asp Tyr Phe Ala Glu Gly  
 435 440 445

ATA AAG GGT TGATTAATCA TTTAATCATG TATTATGAAG TTGGATGAAA 1392  
 Ile Lys Gly  
 450

TCCTCTGTTT CATCTCTATT GTTTAAACAA TAATTTTTTT CCATTGAACT TTTTGGAGTC 1452  
 AATAAAAAAA AAAAAAAAATG AAAAAAACTCA GTTATTTTTT TTTTTTTTTT 1512  
 25 TTTTTTTTTT 1521

SEQ ID NO: 7  
 SEQUENCE LENGTH: 10  
 SEQUENCE TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 Arg Phe Leu Gly Ile Thr Gly Ser Pro Lys  
 1 5 10

SEQ ID NO: 8  
 SEQUENCE LENGTH: 8  
 SEQUENCE TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 Ile His Met Asp Ala Phe Ala Lys  
 1 5

SEQ ID NO: 9

SEQUENCE LENGTH: 10  
SEQUENCE TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

Gly Val Glu Ile Gly Val Ser Leu Pro Lys  
1 5 10

SEQ ID NO: 10

SEQUENCE LENGTH: 8  
SEQUENCE TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL SEQUENCE: No  
SEQUENCE DESCRIPTION

Ala Ser Leu Ser Leu Thr Leu Lys  
1 5

SEQ ID NO: 11

SEQUENCE LENGTH: 14  
SEQUENCE TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL SEQUENCE: No  
SEQUENCE DESCRIPTION

His Tyr Val Pro Leu Ser Gly Asn Leu Leu Met Pro Ile Lys  
1 5 10

SEQ ID NO: 12  
SEQUENCE LENGTH: 14  
SEQUENCE TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL SEQUENCE: No  
SEQUENCE DESCRIPTION

Val Arg Ala Thr Tyr Val Leu Ser Leu Ala Glu Ile Gln Lys  
1 5 10

SEQUENCE DESCRIPTION

Val Arg Ala Thr Tyr Val Leu Ser Leu Ala Glu Ile Gln Lys  
1 5 10

SEQ ID NO: 13  
SEQUENCE LENGTH: 8

SEQUENCE TYPE: amino acid  
 TOPOLOGY: linear  
 5 MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 Ile His Met Asp Ala Phe Ala Lys  
 10 1 5  
 SEQ ID NO: 14  
 SEQUENCE LENGTH: 9  
 SEQUENCE TYPE: amino acid  
 15 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 20 Lys Ile His Met Asp Ala Phe Ala Lys  
 1 5  
 SEQ ID NO: 15  
 SEQUENCE LENGTH: 8  
 SEQUENCE TYPE: amino acid  
 TOPOLOGY: linear  
 30 MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 Lys Ile His Met Asp Ala Phe Ala  
 35 1 5  
 SEQ ID NO: 16  
 SEQUENCE LENGTH: 23  
 SEQUENCE TYPE: nucleic acid  
 40 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 45 AARATHCAYA TGGAYGCITT YGC  
 SEQ ID NO: 17  
 SEQUENCE LENGTH: 23  
 SEQUENCE TYPE: nucleic acid  
 50  
 55

23

STRANDEDNESS: single  
 TOPOLOGY: linear  
 5 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 CTCGAGTTTT TTTTTTTTT TTT 23  
 10 SEQ ID NO: 18  
 SEQUENCE LENGTH: 26  
 SEQUENCE TYPE: nucleic acid  
 15 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 20 SEQUENCE DESCRIPTION  
 TTCACCATGG AGCAAATCCA AATGGT 26  
 SEQ ID NO: 19  
 SEQUENCE LENGTH: 17  
 25 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 30 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 CGAGTCGCCC TCATCAC 17  
 35 SEQ ID NO: 20  
 SEQUENCE LENGTH: 16  
 SEQUENCE TYPE: nucleic acid  
 40 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 45 SEQUENCE DESCRIPTION  
 AACAGCTATG ACCATG 16  
 SEQ ID NO: 21  
 SEQUENCE LENGTH: 6  
 50 SEQUENCE TYPE: amino acid  
 TOPOLOGY: linear

55



MOLECULE TYPE: peptide  
 HYPOTHETICAL SEQUENCE: No  
 5 SEQUENCE DESCRIPTION  
 Asp Phe Gly Trp Gly Lys  
 1 5  
 SEQ ID NO: 22  
 10 SEQUENCE LENGTH: 17  
 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: single  
 15 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 20 GAYTTYGGIT GGGGIAA 17  
 SEQ ID NO: 23  
 SEQUENCE LENGTH: 21  
 25 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 30 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 TGGCAACTGT CTTGCGTCAT G 21  
 35 SEQ ID NO: 24  
 SEQUENCE LENGTH: 23  
 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: single  
 40 TOPOLOGY: linear  
 MOLECULE TYPE: synthetic DNA  
 HYPOTHETICAL SEQUENCE: No  
 SEQUENCE DESCRIPTION  
 45 CCATGTCAGG TGTGAGGTTT AAC 23  
 SEQ ID NO: 25  
 SEQUENCE LENGTH: 20  
 50 SEQUENCE TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 55

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

ATCGTTTCGC ATGATTGAAC

20

SEQ ID NO: 26

SEQUENCE LENGTH: 20

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

TCAGAAGAAC TCGTCAAGAA

20

SEQ ID NO: 27

SEQUENCE LENGTH: 53

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

GGGATCCAAC A ATG GAG CAA ATC CAA ATG GTG GCC GTG ATC GAA ACG TGT 50

Met Glu Gln Ile Gln Met Val Ala Val Ile Glu Thr Cys

1

5

10

AGA

53

Arg

15

SEQ ID NO: 28

SEQUENCE LENGTH: 16

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

GTAACACGAC GGCCAT

16

SEQ ID NO: 29

SEQUENCE LENGTH: 45

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

GGGATCCAAC A ATG GAG CAA ATC CAA ATG GTG AAC ATT CTC GAA C

45

Met Glu Gln Ile Gln Met Val Asn Ile Leu Glu

1

5

10

SEQ ID NO: 30

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

CTCGGAGGAA TTCGGCACGA C

21

SEQ ID NO: 31

SEQUENCE LENGTH: 35

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: synthetic DNA

HYPOTHETICAL SEQUENCE: No

SEQUENCE DESCRIPTION

AGTCGGATCC AACA ATG ACC ACC CTC CTC GAA TCC

35

Thr Thr Leu Leu Glu Ser

1

5

#### Claims

1. Gene encoding a protein having aromatic acyl group transfer activity or a derivative thereof having said enzymatic activity.
2. The gene according to claim 1 obtainable by cloning using as the primer a nucleotide sequence encoding the amino acid sequence as set forth in SEQ ID No. 21.
3. The gene according to claim 2 wherein said primer has the nucleotide sequence as set forth in SEQ ID No. 22.

4. The gene according to claim 1 or 2 encoding any of the amino acid sequences as set forth in SEQ ID No. 1 to 6, or a modified amino acid sequence in which the above amino acid sequence has been modified by addition or removal of one or more amino acids, or substitution with other amino acid(s).
5. The gene according to claim 1 or 2 encoding a protein, which gene is capable of hybridizing with part or all of the nucleotide sequence encoding any of the amino acid sequences as set forth in SEQ ID No. 1 to 6 under the condition of 5 x SSC and 50°C, and which protein has aromatic acyl group transfer activity.
6. The gene according to claim 1 or 2 encoding a protein, which gene is capable of hybridizing with part or all of the nucleotide sequence encoding any of the amino acid sequences as set forth in SEQ ID No. 1 to 6 under the condition of 2 x SSC and 50°C and which protein has aromatic acyl group transfer activity.
7. The gene according to claim 1 or 2 encoding a protein which has an amino acid sequence having a homology of at least 15% or higher with any of the amino acid sequences as set forth in SEQ ID No. 1 to 6, and which has aromatic acyl group transfer activity.
8. The gene according to claim 1 or 2 encoding a protein which has an amino acid sequence having a homology of at least 30% or higher with any of the amino acid sequences as set forth in SEQ ID No. 1 to 6, and which has aromatic acyl group transfer activity.
9. A vector comprising a gene according to any of claims 1 to 8.
10. A host transformed with a vector according to claim 9.
11. A host according to claim 10 wherein said host is a microbial or animal cell.
12. A host according to claim 10 wherein said host is a plant cell or a plant body.
13. Protein encoded by a gene according to any of claims 1 to 8.
14. Protein having aromatic acyl group transfer activity which is obtainable by affinity chromatography of a crude enzyme extract of a plant with a resin on which Cibacron Blue 3GA is immobilized.
15. Protein capable of specifically binding to antibody to the protein of claim 13 or 14 and which has aromatic acyl group transfer activity.
16. A method for producing a protein having aromatic acyl group transfer activity wherein a host according to claim 10 is cultured or grown, and then a protein having aromatic acyl group transfer activity is recovered from said host.
17. A method for producing a protein having aromatic acyl group transfer activity wherein a crude enzyme extract of a plant is treated with an affinity chromatography using a resin on which Cibacron Blue 3GA is immobilized.
18. A method for producing a protein having aromatic acyl group transfer activity, comprising the step of specifically binding to antibody to the protein of any of claims 13 to 15.
19. A method for acylating a pigment, comprising bringing said pigment to the action of the protein of any of claims 13 to 15.
20. A method for acylating a pigment in a plant, comprising the steps of introducing a gene according to any one of claims 1 to 8 into the plant, allowing said gene to express, and acylating the pigment in the plant with the protein produced.
21. A method for stabilizing a pigment wherein said pigment is acylated by the action of the protein according to any of claims 13 to 15.
22. A method for stabilizing a pigment in a plant, comprising the steps of introducing the gene according to any of claims 1 to 8 into a plant, allowing said gene to express, and acylating the pigment in the plant with the protein produced.

**EP 0 810 287 A1**

23. A method for controlling the color of flowers, comprising the steps of introducing the gene according to any of claims 1 to 8 into a plant, allowing said gene to express, and acylating the pigment in the plant with the protein produced.

5 24. The method according to any of claims 19 to 23 wherein the pigment is anthocyanin.

25. A plant whose color has been controlled by introducing therinto a gene according to any of claims 1 to 8, or its progeny having the same property, or tissues thereof.

10 26. The plant tissue according to claim 25 wherein said tissue is a flower.

27. A cut flower of the plant according to claim 25 or its progeny having the same property.

15

20

25

30

35

40

45

50

55

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP96/00348

## A. CLASSIFICATION OF SUBJECT MATTER

Int. C1<sup>6</sup> C12N15/54, C12N9/10, C12N1/21, C12N5/00, A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. C1<sup>6</sup> C12N15/54, C12N9/10, C12N1/21, C12N5/00, A01H5/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS ONLINE, BIOSIS, WPI/WPIL, GENETYX

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Phytochemistry (Oxford) Vol. 35(6) 1994 Javelle F. et al. "Formation of omega-feruloyloxypalmitic acid by an enzyme from wound-healing potato tuber discs" p. 1419-1424	1 - 27
A	Phytochemistry (Oxford) Vol. 32(1) 1993 Murakoshi I. et al. "Acyltransferase for lupin alkaloids in lupinus-hirsutus" p. 87-91	1 - 27
A	Plant Physiol (BETHESDA) Vol. 98(4) 1992 Javelle F. et al. "Purification and properties of putrescine hydroxycinnamoyl transferase from tobacco nicotiana-tabacum cell suspensions" p. 1264-1269	1 - 27
A	Phytochemistry(Oxford) Vol. 31(3) 1992 Macheix J-J et al. "Partial purification and characterization of hydroxycinnamoyl coA transferase from apple and date fruits" p. 767-772	1 - 27

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"Z" document member of the same patent family

Date of the actual completion of the international search

May 14, 1996 (14. 05. 96)

Date of mailing of the international search report

May 21, 1996 (21. 05. 96)

Name and mailing address of the ISA/

Japanese Patent Office

Facsimile No.

Authorized officer

Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP96/00348

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Phytochemistry(Oxford) 30(8) 1991 Negrel J. et al. "Tyramine hydroxycinnamoyl transferase in the roots of wheat and barley seedlings" p. 2519-2522	1 - 27
A	Phytochemistry(Oxford) Vol. 30(5) 1991 Witte L. et al. "Quinolizidine alkaloids and the enzymatic syntheses of their cinnamic and hydroxycinnamic acid esters in lupinus-angustifolius and lupinus-luteus" p. 1493-1498	1 - 27
A	Phytochemistry(Oxford) Vol. 22(11) 1983 "Agmatine coumaroyl transferase EC-2.3.1. - from barley hordeum-vulgare seedlings" p. 2401-2404	1 - 27
A	Phytochemistry Vol. 19(8) 1980 Zenk M. H. et al. "Partial purification and properties of hydroxy cinnamoyl transferase from higher plants" p. 1625-1630	1 - 27
A	Nature Vol. 366, 1993 Edwina C. Corinsh et al. "Cloning and expression of cytochrome P450 genes controlling flower colour" p. 276-279	1 - 27
A	JP, A, 6-500239 (Int Flower dev. pty. Ltd.), January 31, 1994 (13. 01. 94) & EP, A, 522880 & US, A, 5349125	1 - 27

Form PCT/ISA/210 (continuation of second sheet) (July 1992)